

TECHNISCHE UNIVERSITÄT BERLIN
FAKULTÄT II - MATHEMATIK UND NATURWISSENSCHAFTEN
INSTITUT FÜR MATHEMATIK

MASTERARBEIT

IM STUDIENGANG TECHNOMATHEMATIK

Entwicklung von parallelisierbaren Gradienten-basierten Verfahren zur automatisierten, Ersatzmodell-gestützten Optimierung unter Nebenbedingungen für CFD-FEM-Verdichterdesign

Franziska Krüger

betreut von Prof. Dr. Fredi Tröltzsch

externe Betreuer: Dr. Achim Basermann (DLR)
und Dr. Christian Voß (DLR)

4. Juni 2012

Die selbstständige und eigenhändige Ausfertigung versichert an Eides statt

Berlin, den 4. Juni 2012

(Franziska Krüger)

Vorwort und Danksagung

Die vorliegende Masterarbeit ist im Rahmen meiner Tätigkeit als studentische Mitarbeiterin in der Einrichtung „Simulations- und Softwaretechnik“ des Deutschen Zentrums für Luft- und Raumfahrt (DLR) entstanden. Ich war in der Gruppe „High Performance Computing“ der Abteilung „Verteilte Systeme und Komponentensoftware“ beschäftigt. Die Arbeit wurde in Zusammenarbeit mit dem DLR-Institut für Antriebstechnik, Abteilung „Fan und Verdichter“ entwickelt. Sie bildet den Abschluss meines Studiums der Technomathematik an der Technischen Universität Berlin.

Mein erster Dank gilt Herrn Professor Dr. Fredi Tröltzsch für die fachliche Betreuung seitens der Technischen Universität Berlin. Weiter danke ich Herrn Andreas Schreiber für die Möglichkeit, diese Masterarbeit im DLR anzufertigen.

Besonderer Dank geht an Herrn Dr. Achim Basermann und Herrn Dr. Christian Voß, meine beiden fachlichen Betreuer im DLR. Ich danke ihnen für die vielseitige Unterstützung, die Geduld und ihr außerordentliches Engagement.

Für die Beantwortung meiner Fragen und die regen Diskussionen zum Thema Kriging danke ich Marcel Aulich (DLR). Weiterhin danke ich Melven Zöllner (DLR), der mir insbesondere in der Anfangsphase der Programmierung in C stets hilfreich zur Seite stand.

Schließlich bedanke ich mich bei meiner Familie und meinen Freunden, die mir durch ihre Unterstützung geholfen haben, diese Arbeit zu erstellen.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Einführung in die Grundlagen der Formoptimierung	1
1.2	Herangehensweise des DLR	2
1.3	Aufgabenstellung und Gliederung der Arbeit	3
2	Der automatisierte Optimierer AutoOpti	5
2.1	Der Verdichter und seine Anwendung	5
2.2	Das Programm AutoOpti	6
2.2.1	Programmaufbau und Algorithmus	6
2.2.2	Design-Prozesskette	9
3	Das Interpolationsverfahren Kriging	13
3.1	Allgemeines zur Interpolation	13
3.1.1	Nichtstatistische Interpolationsverfahren	13
3.1.2	Statistische Interpolationsverfahren	15
3.2	Das Kriging-Verfahren	15
3.2.1	Variogramm	16
3.2.2	Ansatz des Kriging-Verfahrens	22
3.2.3	Ordinary Kriging	24
3.2.4	Weitere Kriging-Methoden	28
3.2.5	Gradient Enhanced Kriging	29
4	Nichtlineare Optimierung unter Nebenbedingungen	35
4.1	Logarithmische Barriere-Methode	35
4.2	Augmented-Lagrange-Verfahren	37
4.3	Verfahren zur unrestringierten Optimierung	39
4.3.1	Gradientenverfahren	39
4.3.2	Fletcher-Reeves-Verfahren	41
4.3.3	Polak-Ribière-Verfahren	42
4.3.4	Quasi-Newton-Verfahren	42
4.3.5	Schrittweitenbestimmung mit dem Armijo-Verfahren	45
4.4	Parallelisierungsaspekte	46
4.4.1	Parallelisierbarkeit der Schrittweitensteuerung	47
4.4.2	Parallelisierung in AutoOpti und TRACE	48
5	Auswertung der implementierten Verfahren	51
5.1	Diskussion anhand einer Testfunktion	51
5.1.1	Optimierung bei eindeutig bestimmtem Minimum	53
5.1.2	Optimierung bei verschiedenen Startpunkten	57
5.1.3	Optimierung bei Restriktion der Unsicherheit	59
5.1.4	Ergebniszusammenfassung der Auswertung der Testfunktion	68
5.2	Diskussion anhand eines Verdichterdesignbeispiels	69
5.2.1	Optimierung der gewichteten Summe ohne Restriktion	72
5.2.2	Optimierung unter Restriktion des Totaldruckverlustes	80
5.2.3	Optimierung der gewichteten Summe mit Restriktion	85
5.2.4	Zusammenfassung der Verdichterdesign-Ergebnisse	89

6	Fazit und Ausblick	93
A	Begriffserklärungen	95
A.1	Einführung in die Mehrzieloptimierung	95
A.2	Begriffe zur Parallelisierung	96
B	Implementierung der Verfahren	99
B.1	Eingabedateien	99
B.2	Interface zwischen Optimierer und Direct GEK	102
	Literaturverzeichnis	103
	Abbildungsverzeichnis	107

1 Einleitung

Die vorliegende Masterarbeit behandelt aerodynamisches Verdichterdesign, also die Optimierung der Form von Verdichtern und sogenannten Verdichterprofilen (Profile von Verdichterschaufeln, siehe Abb. 1.1). Die Formoptimierung (engl. *shape optimization*) kann als ein Teilgebiet der Theorie optimaler Steuerungen angesehen werden.

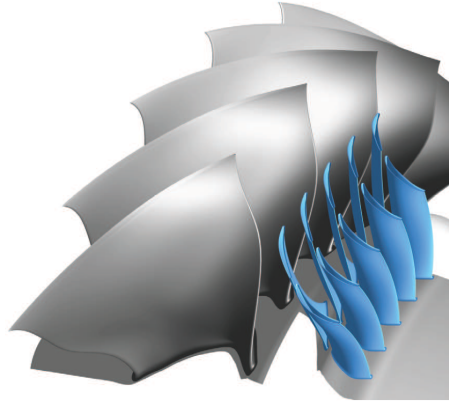


Abbildung 1.1: Schaufeln eines Verdichters. (Quelle: [1])

Im folgenden Abschnitt 1.1 werden zunächst die theoretischen Grundlagen der Formoptimierung skizziert. Anschließend wird auf die Herausforderungen eingegangen, die sich dem DLR bei der Umsetzung dieser theoretischen Herangehensweise stellen (vgl. Abschnitt 1.2), um schließlich daraus die Aufgabenstellung dieser Masterarbeit abzuleiten (vgl. Abschnitt 1.3).

1.1 Einführung in die Grundlagen der Formoptimierung

Die zu optimierende Form wird durch eine gewisse Anzahl geometrischer Variablen beschrieben. Im Verdichterprofildesign handelt es sich bei diesen geometrischen Variablen unter anderem um den Staffelungswinkel β_{ST} , die Länge L oder die maximale Dicke d des Profils (vgl. Abb. 1.2). Die zu optimierende Zielfunktion ist beispielsweise der Totaldruck p_T , welcher maximiert werden soll.

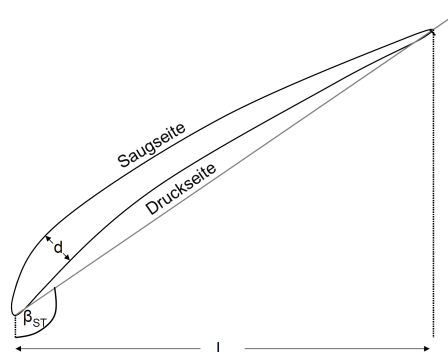


Abbildung 1.2: Skizze eines Verdichterschaufelprofils mit geometrischen Variablen.

Die übliche Vorgehensweise zur Lösung von Formoptimierungsproblemen kann in drei Schritte unterteilt werden. Bei gegebener Form wird zunächst das Strömungsfeld mit Hilfe eines geeigneten Strömungslösers berechnet. Im nächsten Schritt werden die Formableitungen (engl. *shape derivatives*) ermittelt, also die Ableitungen der zu optimierenden Zielfunktion bezüglich der geometrischen Variablen. Dies erfolgt meist durch Lösung der adjungierten Gleichungen. Abhängig von den so berechneten Gradienten und der Wahl des Gradienten-basierten Optimierungsverfahrens wird schließlich die Form aktualisiert (vgl. [2]). Diese Schritte werden iterativ bis zur Konvergenz gegen eine lokal optimale Form fortgeführt.

1.2 Herangehensweise des DLR

Die Herausforderungen, denen das Institut für Antriebstechnik des DLR bei der in Abschnitt 1.1 skizzierten Herangehensweise gegenübersteht, sind vielfältig. Zwar existiert der unter anderem am DLR entwickelte Strömungslöser *TRACE* zur Lösung der Reynolds-Averaged-Navier-Stokes-Gleichung (RANS-Gleichung) schon länger, allerdings war die Ermittlung der Formableitungen und somit eine Formoptimierung auf die zuvor dargelegte Weise lange Zeit problematisch. Erst seit Kurzem ist die Berechnung der Formableitungen durch die Entwicklung des Programmpakets *adjointTRACE* zur Lösung der diskret adjungierten Reynolds-Gleichungen möglich.

Ein weiteres Problem stellt der enorme Zeitbedarf des zuvor beschriebenen Verfahrens zur Formoptimierung dar. Jede Iteration dieses Gradienten-basierten Optimierungsverfahrens erfordert die Lösung der diskretisierten RANS-Gleichung sowie der diskret adjungierten Reynolds-Gleichungen.

Schließlich stellt das Verdichterdesign meist einen Kompromiss zwischen verschiedenen Anforderungen dar. So soll beispielsweise der Totaldruck maximiert sowie der Totaldruckverlust und der Drall der Strömung nach Austritt aus dem Verdichter minimiert werden. Aus diesem Grund legt das Institut für Antriebstechnik des DLR bei Verfahren zur Formoptimierung großen Wert auf die Möglichkeit zur Optimierung mehrerer Zielfunktionen (sog. Vektor- oder auch Mehrzieloptimierung, engl. *multi-objective optimization*, nähere Erläuterungen siehe Anhang A.1). Erst nach Reduzierung der Mehrzieloptimierungsaufgabe auf ein Optimierungsproblem mit einer reellwertigen Zielfunktion ist eine Lösung mit der oben dargestellten Herangehensweise möglich.

Aus den genannten Gründen wurde am Institut für Antriebstechnik des DLR das Programmpaket *AutoOpti* zur Verdichteroptimierung entwickelt. Dieses basiert auf genetischen Algorithmen. Die Evolutionsstrategien eignen sich hervorragend zur Lösung von Mehrzieloptimierungsproblemen, da diese Algorithmen die Mehrzielstruktur erhalten und sogenannte Pareto-Fronten (vgl. Anhang A.1) generieren können. Aufgrund des Evolutionsansatzes werden keine Formableitungen benötigt und die Lösung der adjungierten Gleichungen ist somit nicht notwendig. Ein Nachteil der genetischen Algorithmen ist ihre langsame Konvergenz, insbesondere da jede Funktionsauswertung von neu generierten Verdichterkonfigurationen die zeitaufwendige Lösung der diskretisierten RANS-Gleichung erfordert.

Zu einer erheblichen Reduzierung des Zeitaufwandes führt hier die parallele Generierung und anschließende Optimierung sogenannter Ersatzmodelle. Diese mit Hilfe des geostatistischen Interpolationsverfahrens Kriging (vgl. Kapitel 3) erstellten Modelle bilden eine Approximation an die zu optimierenden, nicht explizit angebbaren Zielfunktionen und

berechnen sogleich die Unsicherheiten dieser approximierten Zielfunktionswerte. Bisher erfolgt die Ersatzmodelloptimierung gleichermaßen mit genetischen Algorithmen.

1.3 Aufgabenstellung und Gliederung der Arbeit

Ziel dieser Arbeit ist nun die Entwicklung und Implementierung parallelisierbarer Gradienten-basierter Verfahren zur Optimierung der Ersatzmodelle. Die hierfür benötigten Gradienten der Ersatzmodelle werden durch adjointTRACE und ein spezielles Kriging-Verfahren, das Direct Gradient Enhanced Kriging (Direct GEK, vgl. Abschnitt 3.2.5), zur Verfügung gestellt. Da Restriktionen der geometrischen Variablen, der Unsicherheiten und eventuell weiterer aerodynamischer Größen zu beachten sind, betrachten wir Verfahren zur Optimierung nichtlinearer Zielfunktionen unter nichtlinearen Nebenbedingungen. Aufgrund der Gradienten-basierten Verfahren ist eine direkte Optimierung mehrerer Zielfunktionen nun nicht mehr möglich. Abschnitt 5.2 zeigt verschiedene Lösungsansätze für dieses Problem.

Die Arbeit ist wie folgt gegliedert. In Kapitel 2 werden Verdichter beschrieben und ihre Anwendungen im DLR erläutert. Darüber hinaus geben wir mit Einführung in das Programmpaket AutoOpti einen Einblick in den bisherigen Stand der Verdichteroptimierung am Institut für Antriebstechnik des DLR. Zudem werden geplante und in Zukunft mögliche Einsatzgebiete des im Rahmen dieser Arbeit implementierten Optimierers als Teil des Programmpakets AutoOpti dargelegt. Es wird herausgestellt, wo und inwieweit das Kriging-Verfahren zur Erstellung der Ersatzmodelle zum Einsatz kommt und warum eine Optimierung der Ersatzmodelle sinnvoll ist. Des Weiteren werden die Vor- und Nachteile der in dieser Arbeit untersuchten Ersatzmodelloptimierung mittels Gradienten-basierter Verfahren dargelegt.

Zum besseren Verständnis wird in Kapitel 3 genauer auf die mathematischen Hintergründe des Kriging-Verfahrens eingegangen, welches die im Rahmen dieser Arbeit zu optimierenden Ersatzmodelle erstellt. Da wir zum Thema Kriging keine passende exakte mathematische Literatur finden konnten, beruhen die Beweise in diesem Kapitel auf eigenen Überlegungen.

Das darauffolgende Kapitel 4 stellt die im Rahmen dieser Arbeit implementierten Verfahren zur Ersatzmodelloptimierung vor. Zur Lösung des nichtlinearen Optimierungsproblems mit nichtlinearen Nebenbedingungen betrachten wir sowohl die logarithmische Barriere-Methode als auch das Augmented-Lagrange-Verfahren. Da beide Methoden auf ein Verfahren zur unrestringierten Minimierung zurückgreifen, stellen wir im Anschluss die unterschiedlichen implementierten Gradienten-basierten Verfahren zur unrestringierten Minimierung vor, wie etwa das Gradientenverfahren oder ein Quasi-Newton-Verfahren. Schließlich werden die Aspekte der Parallelisierbarkeit der implementierten Verfahren analysiert und diskutiert.

Die in Kapitel 4 vorgestellten Verfahren werden in Kapitel 5 anhand zweier Beispiele ausgewertet. Für die Erstellung der zu optimierenden Ersatzmodelle nutzen wir das bereits von Seiten des DLR in AutoOpti implementierte Direct GEK-Programm. Wir betrachten zuerst eine Testfunktion, um im Anschluss ein reales Verdichterdesign-Beispiel zu untersuchen. Dabei wird die Form des Verdichterprofils hinsichtlich des maximalen Totaldruckverhältnisses und minimalen Totaldruckverlustes optimiert. Dieses Mehrzieloptimierungsproblem werden wir in Abschnitt 5.2 durch unterschiedliche Herangehensweisen in ein Problem mit nur einer reellwertigen Zielfunktion umformulieren. Daraufhin werden das

Augmented-Lagrange-Verfahren und die logarithmische Barriere-Methode unter anderem im Hinblick auf die verwendeten Verfahren zur unrestringierten Minimierung verglichen.

2 Der automatisierte Optimierer AutoOpti

AutoOpti ist ein am DLR-Institut für Antriebstechnik in Köln entwickelter automatisierter Optimierer. Ziel von AutoOpti ist die Beschleunigung des aerodynamischen Formoptimierungsprozesses von Verdichtern beziehungsweise Verdichterprofilen (vgl. Abb. 1.1 und Abb. 1.2).

Bevor wir den Optimierungsprozess in AutoOpti und somit den bisherigen Stand der Verdichteroptimierung im DLR näher untersuchen, wird zunächst der Verdichter erklärt und seine Verwendung als Komponente eines Strahltriebwerks erläutert.

2.1 Der Verdichter und seine Anwendung

Verdichter oder auch Kompressoren sind Fluidenergiemaschinen zum Komprimieren von Gasen. Eine spezielle Bauform ist der sogenannte Turboverdichter. Heutzutage existieren zwei Bauarten des Turboverdichters: der Axial- und der Radialverdichter.

Das zu komprimierende Gas strömt beim Axialverdichter parallel zur rotierenden Achse durch den Verdichter und wird meist über mehrere abwechselnd angeordnete rotierende und stationäre Verdichterschaufelreihen, sogenannte Verdichterstufen, komprimiert. In Abbildung 2.1 ist ein solcher mehrstufiger Axialverdichter abgebildet.

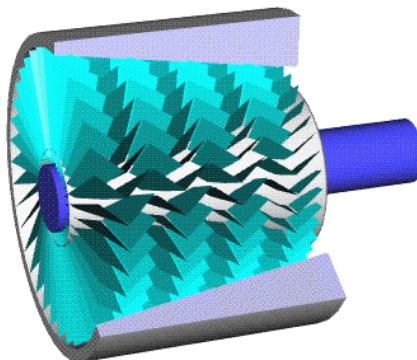


Abbildung 2.1: Mehrstufiger Axialverdichter mit rotierenden Verdichterschaufeln (sog. Rotoren) und stationären Verdichterschaufeln (sog. Statoren), welche abwechselnd angeordnet sind. (Quelle: [3])

Bei einem Radialverdichter wird das Gas dem Zentrum der rotierenden Achse zugeführt und aufgrund der Fliehkräfte radial nach außen abgelenkt. Somit ist bei mehrstufigen Radialverdichtern nach jeder Verdichterstufe eine Umlenkung der Strömung notwendig (vgl. [3, 4]).

Anwendung finden Turboverdichter beispielsweise in Strahltriebwerken. Hierbei handelt es sich in den meisten Fällen um Axialverdichter, welche sich direkt hinter dem Lufteinlauf befinden (siehe Abb. 2.2).

In einem solchen Axialverdichter wird den einströmenden Luftmassen durch die kontinuierlich rotierenden Schaufeln des Rotors kinetische Energie zugeführt, welche aufgrund der Bernoulli-Gleichung (vgl. [5]) in den diffusorförmigen Zwischenräumen der Verdichterschaufeln des Rotors und Stators in Druckenergie umgewandelt wird. Aufgrund des Druckanstiegs verringert sich jedoch die kinetische Energie des Gases, was im Rotor wieder

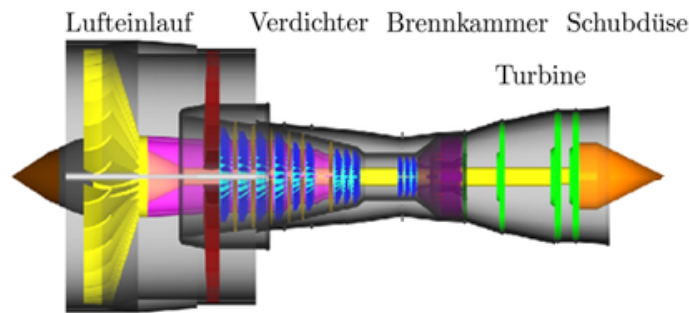


Abbildung 2.2: Skizze eines Strahltriebwerks mit Verdichter und weiteren relevanten Bauteilen. (Quelle: [6])

ausgeglichen wird. Somit besteht eine Axialverdichterstufe meist aus einem Rotor, wo Druck und Geschwindigkeit des Gases zunehmen und einem Stator, wo sich nur der Druck zu Ungunsten der Geschwindigkeit erhöht.

Hinter dem Verdichter befindet sich die Brennkammer, wo das aufgrund der Kompression stark erhitzte Gas mit dem Kraftstoff zusammengeführt und verbrannt wird. Es kommt zu einer Ausdehnung des Gases. Im Anschluss daran trifft das nach hinten austretende Gas auf die Turbine, welche über eine Welle den Verdichter antreibt. Je nach Art des Strahltriebwerks wird der größte Teil der Energie jedoch nicht hier verbraucht, sondern für den Rückstoß in der dahinterliegenden Schubdüse benötigt. Zur Schuberzeugung wird das Gas mit Hilfe der Düse auf eine hohe Geschwindigkeit beschleunigt (vgl. [6]).

2.2 Das Programm AutoOpti

Das von Seiten des DLR implementierte Programmpaket AutoOpti zur Verdichteroptimierung und dessen Algorithmus basieren aus den unter anderem in Abschnitt 1.2 genannten Gründen auf der Evolutionsstrategie. Nach dem Vorbild der Natur (Stichwort *survival of the fittest*) wird die Menge der Individuen, also verschiedene Konfigurationen von Verdichtern bzw. Verdichterprofilen, durch Mutation, Selektion und Vererbung stetig erweitert und verbessert (vgl. [1]).

Da das Verdichterdesign immer einen Kompromiss zwischen verschiedensten Anforderungen wie beispielsweise zu maximierendem Wirkungsgrad, Totaldruckverhältnis und Massenstrom darstellt, wurde bei AutoOpti besonderer Wert auf die Möglichkeit zur gleichzeitigen Optimierung mehrerer Zielfunktionen gelegt (vgl. [1, 7, 8]).

Zudem arbeitet AutoOpti hochparallel und nutzt so alle zur Verfügung stehenden Rechenressourcen. Die genaue Form der Parallelisierung wird unter anderem im folgenden Abschnitt 2.2.1 erläutert.

2.2.1 Programmaufbau und Algorithmus

Den Aufbau und die Struktur des parallelisierten Mehrziel-Evolutionsalgorithmus AutoOpti zeigt Abbildung 2.3. Der *Opti-Master* führt den primären Optimierungsprozess durch. Zur sehr teuren Berechnung des Zielfunktionswertes eines erzeugten Individuums, aufgrund des Evolutionscharakters auch als Fitness bezeichnet, übergibt der Hauptprozess das Individuum an einen *Opti-Slave*. Dieser durchläuft zur Fitnessberechnung die Design-

Da sich im Fall von AutoOpti aufgrund der CFD- und CSM-Simulation ein enormer numerischer Aufwand hinter der Prozesskette des Designs verbirgt, wurde viel Wert darauf gelegt, dass möglichst wenige Individuen bzw. Verdichterkonfigurationen diese Prozesskette durchlaufen müssen. Um dies zu ermöglichen, sind einige Erweiterungen in AutoOpti integriert, von denen wir im Folgenden nur die wichtigste nennen wollen.

Ein großer Nachteil von Evolutionsalgorithmen ist deren langsame Konvergenz. Vor allem für Probleme mit teurer Zielfunktionsauswertung, wie dies im Verdichterdesign aufgrund der Design-Prozesskette der Fall ist, resultiert daraus eine lange Rechenzeit. Zur Beschleunigung und Verbesserung des Optimierungsprozesses werden daher in AutoOpti einige Approximationsmodelle, sogenannte Ersatz- oder auch Metamodelle, verwendet (siehe Abb. 2.3, links). Während der ursprüngliche Optimierungsprozess (Abb. 2.3, rechte Seite) arbeitet, läuft parallel dazu ein zweiter Prozess (linke Seite). Dieser beschäftigt sich mit dem Erstellen von Ersatzmodellen, dem sogenannten Training, und der anschließenden Optimierung dieser Ersatzmodelle. Beide Prozesse kommunizieren über die Datenbank und das Interface.

Das Ersatzmodell wird mit Hilfe von Stützpunkten, d.h. vorher berechneten Individuen bzw. Verdichterkonfigurationen aus der Datenbank *DataBase* erstellt und dient dazu, den Fitness- bzw. Zielfunktionswert eines beliebigen Individuums respektive einer beliebigen Verdichterkonfiguration zu approximieren. In AutoOpti werden diese Ersatzmodelle mit Hilfe des Interpolationsverfahrens Kriging sowie durch Neuronale Netze (siehe [10]) berechnet. Im Rahmen dieser Arbeit wird zum Erstellen der Ersatzmodelle das von Seiten des DLR in AutoOpti implementierte Kriging-Programm genutzt. Auf die mathematischen Hintergründe des Kriging-Verfahrens und das in AutoOpti verwendete Direct GEK gehen wir im folgenden Kapitel 3 näher ein.

Die mit Hilfe der Optimierung der Ersatzmodelle berechneten Individuen können nun über das Interface in den ursprünglichen Optimierungsprozess integriert werden (vgl. Abb. 2.3).

Bisher erfolgte die Optimierung der Ersatzmodelle wiederum mit einer evolutionären Strategie. Ziel dieser Arbeit ist es, statt genetischer Algorithmen ein Gradienten-basiertes Verfahren zur Optimierung der Ersatzmodelle zu implementieren. Diese zeigen im Vergleich zu Evolutionsalgorithmen üblicherweise ein schnelleres Konvergenzverhalten. Aufgrund der Nichtlinearität des Ersatzmodells und eventuell zu beachtender nichtlinearer Nebenbedingungen, wie etwa einer Restriktion der Unsicherheit oder anderer aerodynamischer Größen, betrachten wir im Folgenden nichtlineare Optimierungsalgorithmen mit nichtlinearen Nebenbedingungen. Auf die hierfür im Rahmen der Arbeit implementierten Verfahren wird in Kapitel 4 näher eingegangen.

Da Gradienten-basierte Verfahren nur eine Zielfunktion optimieren können, wird die Betrachtung mehrerer Zielfunktionen durch diese Herangehensweise erschwert. In Abschnitt 5.2 werden unterschiedliche Lösungsansätze für dieses Problem aufgezeigt. Ein weiterer Nachteil Gradienten-basierter Optimierungsverfahren ist deren Konvergenz gegen lokale und nicht globale Minima.

Eine Gradienten-basierte Formoptimierung im Sinne der in der Einleitung dargestellten Herangehensweise (vgl. Abschnitt 1.1) ist mit Hilfe des im Rahmen dieser Arbeit programmierten Optimierers zukünftig ebenso möglich, da die nötigen Gradienteninformationen beziehungsweise Formableitungen durch den Löser adjointTRACE für die diskret adjungierten Reynolds-Gleichungen zur Verfügung stehen (vgl. [11]). Somit besteht die Möglichkeit, dass die im Rahmen dieser Arbeit implementierten Optimierungsalgorithmen nicht nur auf den Ersatzmodellen, sondern auch auf den normalen Modellen Anwendung finden.

2.2.2 Design-Prozesskette

Die Prozesskette des Designs wird bei jeder Auswertung der Zielfunktion aufgerufen und vollständig durchlaufen. Zur besseren Einschätzung der Auswirkungen einer solchen teuren Zielfunktionsauswertung bzw. Fitnessberechnung wird die Design-Prozesskette im Folgenden kurz vorgestellt. Sie gliedert sich in die in Abbildung 2.3, rechts, dargestellten verschiedenen Abschnitte. Zeitaufwändige Prozesse sind hier insbesondere die CFD-Strömungssimulation mittels Finite-Volumen-Methode und die CSM-Strukturanalyse mit Hilfe der Finite-Elemente-Methode.

Skalierung der Parameter

Im ersten Schritt der Design-Prozesskette findet eine Skalierung der Parameter statt. Die aufgrund des Optimierungsprozesses normalisierten Variablen werden wieder auf ihre reale Größe skaliert.

Parametrisierung der Verdichterschaufeln und des Strömungskanals

Anschließend wird die Geometrie des Strömungskanals mit Hilfe einer Interpolation von sogenannten Kontrollpunkten beschrieben. Diese Kontrollpunkte sichern, dass die Geometrie des Luftkanals innerhalb spezifischer Grenzen liegt.

Die Geometrie der Verdichterschaufeln respektive Verdichterprofile wird mit dem im DLR entwickelten Programm *Bladegenerator* generiert. Als Eingabe erhält das Programm ein grobes zweidimensionales Gitter und eine Menge von Profilparametern, die das zweidimensionale Profil der Verdichterschaufeln beschreiben. Zu diesen Parametern zählen beispielsweise der Staffellungswinkel, der Anström- und Abströmwinkel, gewisse Kontrollpunkte auf der Saugseite (Oberseite des Verdichterprofils) zur Beschreibung des Verlaufs der Saugseite und die Dickeverteilung über das gesamte Profil, aus welcher der Verlauf der Druckseite (Unterseite) abgeleitet werden kann (siehe auch Abb. 1.2).

Nach der Generierung der Verdichterprofile in 2D erhalten diese ihre dreidimensionale Form. Zudem wird festgelegt, welche Position die dreidimensionalen Verdichterschaufeln zueinander einnehmen.

Geometrische Restriktionen

In diesem Schritt besteht die Möglichkeit, weitere geometrische Restriktionen an das Profil zu stellen. Diese Einschränkungen können z.B. den Füllfaktor, die maximale Dicke, die maximale Länge der Profillehne und noch viele weitere Eigenschaften der Verdichterschaufel betreffen.

Unter Berücksichtigung dieser zusätzlichen Einschränkungen wird sodann die endgültige 3D-Verdichterschaufel erzeugt.

Gittergenerierung

Eine Erweiterung des kommerziellen Tools *G3DMESH* (vgl. [12]) ist für die automatische Generierung des Gitters zuständig. Das Gitter besteht aus mehreren Blöcken, welche durch die entsprechende Anzahl an Gitterpunkten in der jeweils notwendigen Feinheit unterteilt werden (siehe Abb. 2.4). Beispielsweise muss die Grenzschicht, also die Schicht nahe der

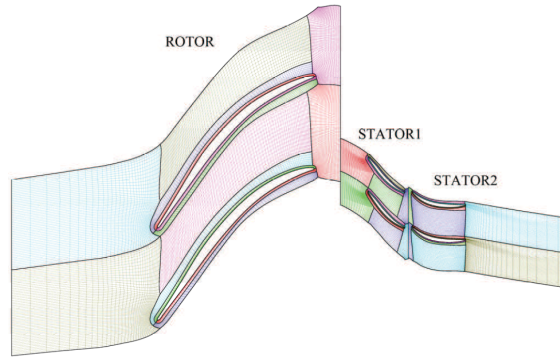


Abbildung 2.4: Beispiel eines Gitters für eine Verdichterstufe mit einem Rotor und zwei Statorn. (Quelle: [1])

Verdichterschaufeln, aufgrund der turbulenten Strömungsvorgänge feiner diskretisiert werden als Strömungsbereiche fern dieser Grenzschicht.

Strömungssimulation mittels Finite-Volumen-Methode

Die Strömungssimulation wird mit TRACE („Turbomachinery Research Aerodynamics Computational Environment“) durchgeführt. TRACE ist ein am DLR-Institut für Antriebstechnik in Kooperation mit MTU Aero Engines entwickeltes CFD-Simulationssystem zur Berechnung von dreidimensionalen Turbinenströmungen (vgl. [1]).

TRACE basiert auf einer Finite-Volumen-Diskretisierung der kompressiblen, instationären Reynolds-Averaged-Navier-Stokes-Gleichung (RANS-Gleichung). In einem rotierenden Bezugssystem lesen sich diese Gleichungen in Erhaltungform wie folgt (vgl. [13]):

$$\begin{cases} \frac{\partial \varrho}{\partial t} + \operatorname{div}(\varrho u) = 0 \\ \frac{\partial(\varrho u)}{\partial t} + \operatorname{div}(\varrho u \otimes u + pI - \tau) + 2\varrho\Omega \times u - \varrho\|\Omega\|^2 r e_r = 0 \\ \frac{\partial(\varrho E)}{\partial t} + \operatorname{div}(\varrho u H_{tot} - \tau u + Q) = 0. \end{cases} \quad (2.1)$$

Hierbei ist ϱ die Dichte, $u = (u_1, u_2, u_3)^T$ der Geschwindigkeitsvektor, p der Druck, τ der viskose Spannungstensor, Ω der Rotationsvektor, r der Radius und e_r der Einheitsvektor in radialer Richtung. Zudem bezeichnet E die Energie, H_{tot} die Staupunkt-Enthalpie und Q die Wärmestromdichte. Des Weiteren sei die Einheitsmatrix mit I , das dyadische Produkt zweier Vektoren mit \otimes und das Kreuzprodukt zweier Vektoren mit \times bezeichnet.

Gleichung (2.1) lässt sich in der allgemeinen Form einer Erhaltungsgleichung

$$\frac{\partial q}{\partial t} + \operatorname{div} F(q) + S(q) = 0 \quad (2.2)$$

schreiben, wobei $q = (\varrho, \varrho u_1, \varrho u_2, \varrho u_3, \varrho E)^T$ der Vektor der Erhaltungsvariablen ist und mit $F(q)$ die Flussfunktion bezeichnet wird. Der Divergenz-Operator ist hier, ähnlich wie in der zweiten Gleichung von (2.1), zeilenweise zu verstehen. Die Coriolis- und Zentrifugalkräfte, die aufgrund des rotierenden Bezugssystems auftreten, werden durch den Quellterm $S(q)$ berücksichtigt.

Die Schließung des Gleichungssystems erfolgt mit dem k - ω -Turbulenzmodell von Wilcox (vgl. [14]).

Die Finite-Volumen-Diskretisierung erfolgt nun auf dem zuvor generierten Gitter. Mit Hilfe des Satzes von Gauß können wir aus Gleichung (2.2) die Finite-Volumen-Formulierung in der i -ten Zelle

$$\frac{\partial}{\partial t}(V_i q_i) + \sum_j F'_{ij}(q) + S_i(q)V_i = 0$$

ableiten. Hierbei sei V_i das Volumen der Zelle. $F'_{ij}(q)$ bezeichnet die Flüsse über alle Ränder j , $S_i(q)$ die Quellen jeder Zelle und q_i die Zellmittelwerte der Erhaltungsvariablen. Zur Berechnung der Flüsse wird das MUSCL-Schema angewandt (vgl. [15]). Dabei handelt es sich um eine Art Upwind-Diskretisierung zur Darstellung der räumlichen Ableitungen von q .

Da ein Verdichter aus vielen, meist identischen Schaufeln besteht, wird die Strömungssimulation mit nur einer oder wenigen Verdichterschaufeln und periodischen Randbedingungen durchgeführt (vgl. [13, 16]).

Strukturanalyse mittels Finite-Elemente-Methode

Die Strukturanalyse oder auch -berechnung dient dem Festigkeitsnachweis der generierten Verdichterschaufeln. Hierbei wird die Verteilung der mechanischen Spannung über die gesamte Schaufel analysiert (vgl. Abb. 2.5). Die Finite-Elemente-Methode ist das gängigste numerische Verfahren um diese Untersuchung durchzuführen.

In AutoOpti wurden zwei Finite-Elemente-Löser zur Strukturanalyse in die Prozesskette eingebunden: das kommerzielle Tool *PERMAS* (siehe [17]) und der Open-Source-Code *CalculiX* (vgl. [18]). Es werden statische und dynamische Strukturanalysen unterstützt.

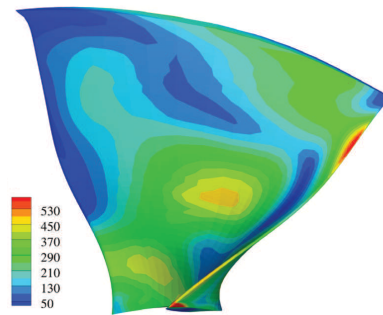


Abbildung 2.5: Verteilung der mechanischen Spannung einer Verdichterschaufel (Berechnung mit CalculiX). (Quelle: [1])

3 Das Interpolationsverfahren Kriging

Die bisherigen Ausführungen und speziell die zeitaufwendige Zielfunktionsauswertung aufgrund der Design-Prozesskette zeigen, dass eine wie in AutoOpti implementierte parallele Optimierung der Ersatzmodelle aus Zeitgründen unbedingt notwendig ist. Das folgende Kapitel gibt einen Einblick in das statistische Interpolationsverfahren Kriging, welches die in AutoOpti und im Rahmen dieser Arbeit betrachteten Ersatzmodelle erstellt. Es wird vorwiegend auf die mathematischen Hintergründe des Kriging-Verfahrens eingegangen.

Wir beginnen mit einer Einführung zum Thema Interpolation und stellen die Vorteile von statistischen Interpolationsverfahren dar. Im Anschluss erläutern wir die Erstellung des Variogramms, welches die Grundlage für das Kriging-Verfahren bildet. Schließlich werden verschiedene Kriging-Verfahren untersucht. Insbesondere das Ordinary Kriging und das Direct Gradient Enhanced Kriging werden aufgrund der Anwendung in dieser Arbeit und in AutoOpti näher analysiert.

3.1 Allgemeines zur Interpolation

Da raumbezogene Daten in der Regel punkthaft erfasst werden, man jedoch meist an einer flächenhaften Darstellung der Messdaten interessiert ist, wird ein Verfahren zur räumlichen Interpolation eingesetzt. Ziel einer Interpolation ist die Approximation eines unbekannten Funktionswertes an einer nicht ausgewerteten Stelle u_0 . Hierbei ist u_0 umgeben von sogenannten Stützstellen u_1 bis u_d mit bekannten Funktionswerten $f(u_i)$, $i = 1, \dots, d$. Im Fall von AutoOpti sind die Stützstellen die schon berechneten Verdichterkonfigurationen in der Datenbasis *DataBase* (vgl. Abb. 2.3). Diese bilden die Grundlage für die Kriging-Interpolation zur Erstellung der Ersatzmodelle, also der Approximation der Zielfunktion.

Man unterscheidet grundsätzlich zwischen nichtstatistischen und statistischen Interpolationsverfahren. Bevor wir die Kriging-Methode, welche in die Kategorie der statistischen Verfahren einzuordnen ist, näher untersuchen, wollen wir einen Blick auf die nichtstatistischen Interpolationsverfahren werfen. Insbesondere sollen die Vorteile eines statistischen Verfahrens gegenüber einem nichtstatistischen Verfahren herausgestellt werden.

3.1.1 Nichtstatistische Interpolationsverfahren

Es gibt eine Vielzahl von Interpolationsverfahren. Alle Verfahren beruhen auf der Annahme, dass eine Ähnlichkeit zwischen räumlich benachbarten Punkten vorliegt. Somit weisen räumlich nah beieinander liegende Punkte eine größere Ähnlichkeit auf als räumlich weit voneinander entfernte Punkte. Ohne diese Annahme wäre eine Interpolation nicht möglich, da aufgrund der dann vorliegenden Unabhängigkeit der Funktionswerte keine Aussage an dem unbeprobten Ort gemacht werden kann.

Aus den Messdaten und der räumlichen Ähnlichkeit wird der Funktionswert an der unbeprobten Stelle geschätzt beziehungsweise approximiert. Die verschiedenen Interpolationsmethoden unterscheiden sich in der Art der Modellierung dieser räumlichen Ähnlichkeit (vgl. [19, 20]).

Im Folgenden wollen wir die Polygon-Methode und die Inverse Distanzgewichtung (engl. *Inverse distance weighting* - IDW) als Beispiele nichtstatistischer Verfahren betrachten.

Bei der Polygon-Methode wird das Untersuchungsgebiet in Polygone aufgeteilt. Wie in Abbildung 3.1 ersichtlich wird, besteht ein Polygon aus all jenen Punkten, die einem beprobten Punkt am nächsten sind. Alle unbeprobten Punkte eines Polygons erhalten dann den Funktionswert des vermessenen Punktes als approximierten Funktionswert. Somit ist die Approximation der wahren Funktion eine Treppenfunktion. An den Polygon-Rändern entstehen folglich Sprungstellen, die oft unerwünscht sind, da sie das reale, meist stetige Verhalten der Beobachtungsvariablen schlecht wiedergeben (vgl. [19]).

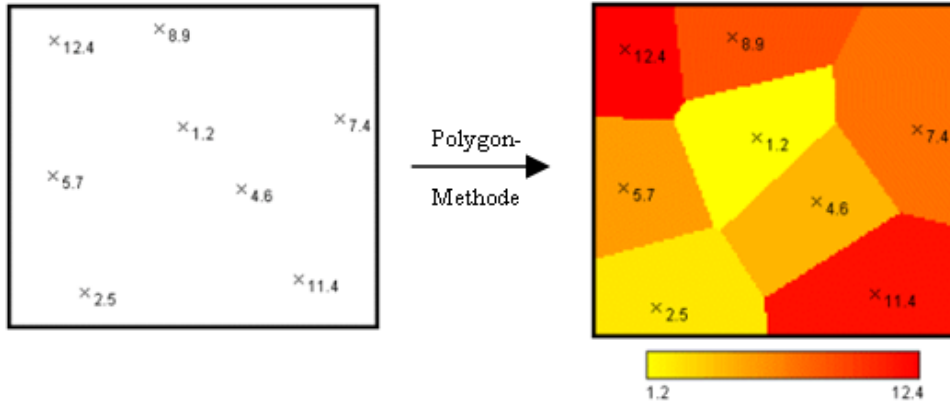


Abbildung 3.1: Interpolation mittels Polygon-Methode. (Quelle: [19])

Bei der Inversen Distanzgewichtung wird der Wert an einem unbeprobten Ort u_0 durch ein gewichtetes Mittel der gemessenen Funktionswerte $f(u_1)$ bis $f(u_d)$ geschätzt. Die Gewichte sind hierbei proportional zum Kehrwert des euklidischen Abstands $d(u_0, u_i)$ zwischen der Stützstelle u_i , $i \in \{1, \dots, d\}$, und dem unbeprobten Ort u_0 (siehe Abb. 3.2). Wir erhalten somit für den gesuchten Funktionswert $f(u_0)$ die Formel

$$f(u_0) = \left(\sum_{i=1}^d \frac{1}{d(u_0, u_i)} \right)^{-1} \cdot \left(\sum_{i=1}^d \frac{1}{d(u_0, u_i)} \cdot f(u_i) \right).$$

Man beachte, dass für u_0 gegen ein u_i , $i \in \{1, \dots, d\}$, im Grenzwert $f(u_0) = f(u_i)$ gilt.

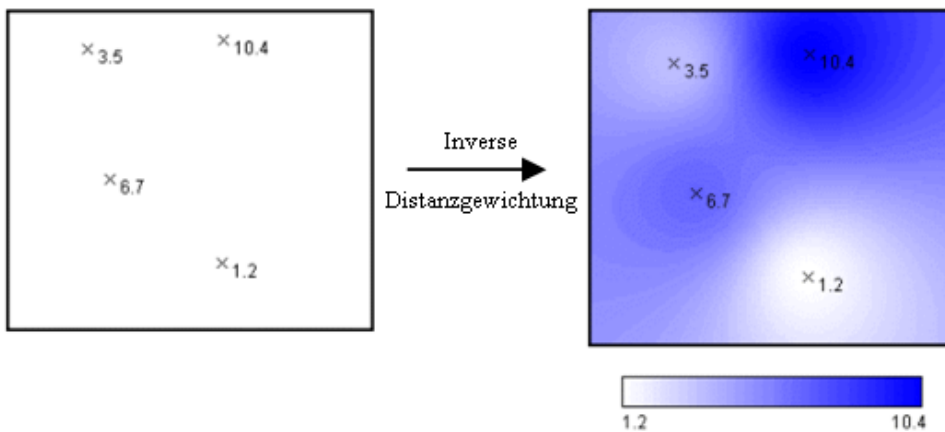


Abbildung 3.2: Interpolation mittels Inverser Distanzgewichtung. (Quelle: [19])

Sobald lokal gehäufte Stützstellen vorliegen, wird ein Nachteil der Inversen Distanzgewichtung deutlich (vgl. [21]). Befinden sich in der Nähe einer Stützstelle weitere Stützstellen, wie in Abbildung 3.3 angedeutet, so sollten diese gehäufte Stützstellen nicht alle volles Gewicht erhalten. Unserer Intuition folgend, würden die Stützstellen u_1 und u_2 stets eine Gewichtung von $1/3$ erhalten, unabhängig von der Anzahl der anderen $d - 2$ Stützstellen um u_3 . Aus diesem Grund sollte die Gewichtung nicht proportional zum Kehrwert des Abstands sein.

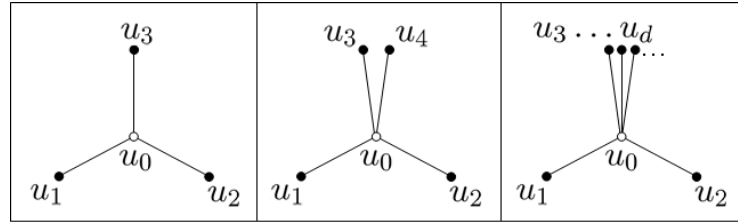


Abbildung 3.3: Inverse Distanzgewichtung bei lokal gehäufte Stützstellen.

Im folgenden Abschnitt werden statistische Interpolationsverfahren vorgestellt und die Vorteile dieser Verfahren gegenüber nichtstatistischen Methoden dargelegt.

3.1.2 Statistische Interpolationsverfahren

Statistische Interpolationsverfahren beruhen auf einem geostatistischen Modell. Dieses Modell gibt den räumlichen Zusammenhang der Punkte im zu untersuchenden Gebiet mit Hilfe von statistischen Größen wie der Varianz und dem Erwartungswert wieder (vgl. Abschnitt 3.2.1).

Das bekannteste statistische Interpolationsverfahren ist das Kriging-Verfahren, welches wir im anschließenden Abschnitt 3.2 näher untersuchen wollen. Wie wir sodann feststellen werden, basiert die Gewichtung der Kriging-Methode nicht nur auf der Entfernung der Stützstellen zum unbeprobten Punkt, sondern sie beinhaltet die gesamte räumliche Verteilung der vermessenen Punkte. Somit lässt sich eine schlechte Wahl der Gewichte aufgrund von lokal gehäufte Stützstellen, wie dies etwa bei der Inversen Distanzgewichtung der Fall ist, vermeiden.

Ein weiterer Vorteil eines statistischen Verfahrens wie der Kriging-Methode ist die Möglichkeit der Angabe eines sogenannten Schätzfehlers, auch als Unsicherheit bezeichnet. Somit lässt sich die Wahrscheinlichkeit bestimmen, mit der der berechnete Funktionswert innerhalb eines bestimmten Toleranzintervalls liegt (vgl. [20]). Dies ist auch bei nichtstatistischen Verfahren möglich, allerdings mit einem weit höheren Aufwand (vgl. [22]).

3.2 Das Kriging-Verfahren

Bei Kriging-Verfahren handelt es sich um eine Gruppe von geostatistischen Methoden, welche die Werte einer Zufallsvariable interpolieren. Basierend auf dem im theoretischen Variogramm (vgl. Abschnitt 3.2.1) dargestellten räumlichen Zusammenhang wird die räumliche Verteilung der Daten geschätzt.

Der Krige-Schätzer ist ein sogenannter BLUE-Schätzer (engl. *best linear unbiased estimator*). Er ist erwartungstreu (engl. *unbiased*), minimiert die Varianz des Schätzfehlers (engl. *best*) und ist zudem als gewichtetes Mittel linear (vgl. [19]).

Benannt wurde das Verfahren nach dem südafrikanischen Bergbauingenieur und Geostatistiker Daniel G. Krige, der als Erfinder des Kriging bezeichnet wird. Auf der Suche nach Goldlagerstätten im Witwatersrand-Höhenzug in Südafrika stellte er fest, dass sich die Vorkommen mit der Methode der kleinsten Quadrate nur schlecht beschreiben lassen (vgl. [23]).

Basierend auf Kriges Masterarbeit [23] entwickelte der französische Mathematiker Georges Matheron die Kriging-Theorie (siehe [24]). Zur gleichen Zeit entwarf Lev S. Gandin in der Sowjetunion dasselbe Verfahren, um es im Bereich der Meteorologie einzusetzen. Gegenwärtig findet Kriging in allen Bereichen der Geowissenschaften sowie in vielen weiteren Forschungsgebieten Anwendung.

Wie schon eingangs erwähnt, existieren mehrere Kriging-Methoden. Zu den klassischen Kriging-Methoden zählen etwa das Simple Kriging, das Ordinary Kriging oder das Universal Kriging. Welche spezielle Methode angewandt wird, ist abhängig von den Eigenschaften der zu interpolierenden Funktion. Die gewählte Methode bestimmt implizit die Art der Berechnung der Kriging-Gewichte und wirkt sich somit auf die Lösung aus.

Im Folgenden definieren wir den Begriff Variogramm und stellen die Eigenschaften eines solchen dar. Im Anschluss wird die Ermittlung des Variogramms mit Hilfe der gemessenen Daten an den Stützstellen u_1 bis u_d erläutert. Der Abschnitt orientiert sich an Streit [19]. Daraufhin betrachten wir einige klassische Kriging-Methoden. Insbesondere studieren wir das Ordinary Kriging (siehe Abschnitt 3.2.3) und das Direct Gradient Enhanced Kriging (vgl. Abschnitt 3.2.5) genauer, da diese Verfahren bei der Erstellung der Ersatzmodelle in AutoOpti Verwendung findet.

3.2.1 Variogramm

Das Variogramm bildet die Grundlage für das statistische Interpolationsverfahren Kriging. Mit Hilfe der gemessenen Daten an den beprobten Stützstellen soll der räumliche Zusammenhang unter Verwendung von stochastischen Begriffen wie Erwartungswert und Varianz bestimmt werden. Da diese Begriffe nur für Zufallsvariablen definiert sind, werden die raumbezogenen Daten als Realisationen von Zufallsvariablen modelliert.

Definition 3.1. *Eine Zufallsvariable ist eine Funktion, die den Ergebnissen eines Zufallsexperiments Werte, sogenannte Realisationen, zuordnet.*

Unsere weiteren Untersuchungen beruhen also auf dem folgenden statistischen Modell: Seien $z(u_1)$ bis $z(u_d)$ die Werte der zu beobachtenden Variable an den paarweise verschiedenen Stützstellen u_1 bis $u_d \in D$ und $D \subseteq \mathbb{R}^n$ das Untersuchungsgebiet[‡] (vgl. Abb. 3.4). Diese werden als Realisationen der Zufallsvariablen $Z(u_1)$ bis $Z(u_d)$ betrachtet. Des Weiteren werden auch für die unbeprobten Orte Zufallsvariablen angenommen. Die Realisationen dieser Zufallsvariablen, d.h. die Werte des Ersatzmodells, werden sodann durch räumliche Interpolation mittels Kriging-Verfahren aus den bekannten Realisationen der Stützstellen geschätzt (siehe Abb. 3.4). Die Menge all dieser Zufallsvariablen $Z = \{Z(u) \mid u \in D\}$ wird als stochastischer Prozess bezeichnet.

Um eine Interpolation vornehmen zu können, ist ein Maß für den räumlichen Zusammenhang zweier ortsbezogener Zufallsvariablen $Z(u)$ und $Z(v)$ des stochastischen Prozesses gesucht. Hierzu verwenden wir die Varianz der Differenz der beiden Zufallsvariablen.

[‡]In unserem Fall sind $z(u_1)$ bis $z(u_d)$ also die berechneten Werte der Zielfunktion an den Stützstellen respektive in der Datenbasis vorhandenen Verdichterkonfigurationen u_1 bis u_d .

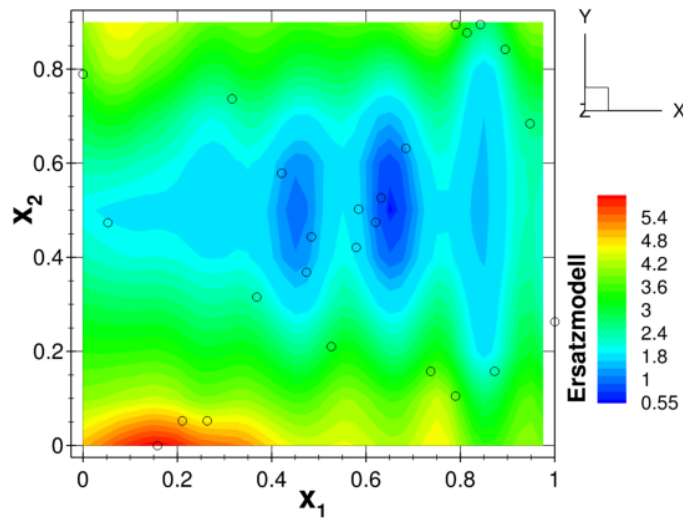


Abbildung 3.4: Beispiel für Stützstellen u_i (schwarze Punkte) im Untersuchungsgebiet D und das auf deren Grundlage mittels Direct Gradient Enhanced Kriging erstellte Ersatzmodell auf D .

Die Varianz ist ein Maß für die mittlere Streuung der Werte einer Zufallsvariablen. Ist die Varianz der Differenz der beiden Zufallsvariablen groß, d.h. streuen die Werte stark, so bedeutet dies, dass die beiden Zufallsvariablen im Mittel sehr unterschiedliche Werte annehmen. Ihr räumlicher Zusammenhang ist somit klein. Ist die Varianz der Differenz $Z(u) - Z(v)$ und somit die Streuung hingegen klein, so schließen wir daraus einen großen räumlichen Zusammenhang. Somit kann $\text{Var}[Z(u) - Z(v)]$ als Maß für den räumlichen Zusammenhang zweier Zufallsvariablen angesehen werden.

Definition 3.2. Sei $D \subseteq \mathbb{R}^n$ das Untersuchungsgebiet und Z ein stochastischer Prozess. Für u und $v \in D$ heißt

$$2\gamma(u, v) = \text{Var}[Z(u) - Z(v)]$$

Variogramm, $\gamma(u, v)$ selbst wird als Semivariogramm bezeichnet (vgl. [25]).

Das Semivariogramm ist somit ein Maß für den räumlichen Zusammenhang zweier Zufallsvariablen und gibt die mittlere Streuung der Differenz dieser an.

Bemerkung 3.1. Es existieren noch weitere Möglichkeiten den räumlichen Zusammenhang zweier Zufallsvariablen wiederzugeben. Hierbei handelt es sich um Methoden, die entweder auf den Korrelationskoeffizienten oder auf die Kovarianz dieser Zufallsvariablen zurückgreifen. Auf diese Methoden soll im Folgenden nicht weiter eingegangen werden. Es sei an dieser Stelle nur erwähnt, dass die drei Methoden eng zusammenhängen. Für weitere Informationen zu diesem Thema verweisen wir auf Isaaks und Srivastava [26].

Die Modellierung des räumlichen Zusammenhangs bei der im Folgenden näher zu betrachtenden Kriging-Methode Ordinary Kriging beruht nun auf der Annahme, dass dieser Zusammenhang nicht von der absoluten geometrischen Lage der Punkte abhängt, sondern nur von der relativen räumlichen Lage der betrachteten Punkte zueinander. Einen ebensolchen stochastischen Prozess bezeichnen wir als intrinsisch oder stationär.

Definition 3.3. Sei $D \subseteq \mathbb{R}^n$ das Untersuchungsgebiet. Ein stochastischer Prozess Z heißt *intrinsisch* oder *stationär*, falls der Erwartungswert $E[Z(u)]$ aller Zufallsvariablen $Z(u)$ des stochastischen Prozesses konstant (aber nicht notwendigerweise bekannt) ist, d.h. $\exists \mu \in \mathbb{R}$, sodass

$$E[Z(u)] = \mu \quad \forall u \in D,$$

und der räumliche Zusammenhang von zwei Zufallsvariablen dieses Prozesses nicht von der absoluten Lage, sondern nur von deren Abstandsvektor $h \in \mathbb{R}^n$ abhängig ist, d.h.

$$\text{Var}[Z(u+h) - Z(u)] = c(h) \quad \forall u, u+h \in D.$$

Somit können wir im Fall eines stationären stochastischen Prozesses das Semivariogramm schreiben als

$$\gamma(h) = \frac{1}{2} \text{Var}[Z(u+h) - Z(u)],$$

wobei sich u und $u+h$ im Untersuchungsgebiet D befinden. Das Semivariogramm ist unabhängig vom Ort $u \in D$, es hängt nur noch vom Abstandsvektor h ab.

Handelt es sich zudem um einen isotropen, also richtungsunabhängigen stochastischen Prozess, so können wir das Semivariogramm weiter zu

$$\gamma(d) = \frac{1}{2} \text{Var}[Z(u+h) - Z(u)]$$

vereinfachen, wobei $d = \|h\|$. In der Regel handelt es sich bei $\gamma(d)$ um eine monoton wachsende Funktion, denn nimmt der räumliche Zusammenhang zweier Zufallsvariablen mit wachsender Distanz d ab, so erhöht sich entsprechend die Varianz der Differenz der beiden Zufallsvariablen (vgl. Abb. 3.5).

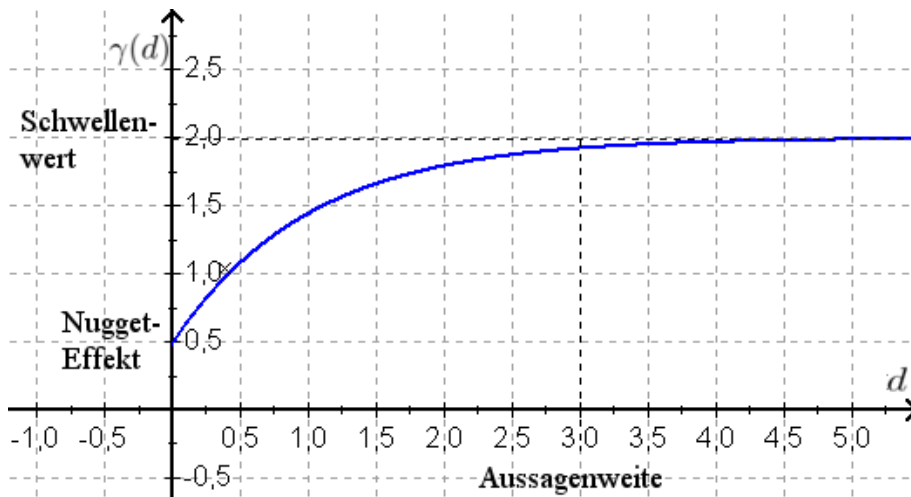


Abbildung 3.5: Beispiel eines Semivariogramms für einen isotropen stationären stochastischen Prozess. Die Kennwerte Aussagenweite, Schwellenwert und Nugget-Effekt sind kenntlich gemacht.

Der Abstand d , bei dem das Semivariogramm einen asymptotischen Schwellenwert (engl. *sill*) erreicht, wird als Aussagenweite (engl. *range*) bezeichnet. Haben zwei Zufallsvariablen einen größeren räumlichen Abstand als die Aussagenweite, so bezeichnet man diese als unabhängig. Verläuft das Semivariogramm nicht durch den Ursprung, so tritt der sogenannte Nugget-Effekt auf. Ein hoher Nugget-Wert (Schnittstelle des Semivariogramms mit der y -Achse) deutet darauf hin, dass nur ein geringer räumlicher Zusammenhang zwischen den Messpunkten vorhanden ist (vgl. [27]).

Die Erstellung des Semivariogramms erfolgt nun in zwei Schritten. Mit Hilfe der aus den Messungen gewonnenen Daten, also den Werten $Z(u_i)$ an den Stützstellen u_1 bis u_d , wird vorerst das sogenannte empirische Semivariogramm berechnet. Im Anschluss wird auf der Grundlage des empirischen Semivariogramms das sogenannte theoretische Semivariogramm modelliert. Dieses werden wir später der Einfachheit halber nur noch als Semivariogramm bezeichnen.

Wir betrachten nun den Fall eines stationären stochastischen Prozesses, d.h. $\exists \mu \in \mathbb{R}$, sodass $E[Z(u)] = \mu$ für alle $u \in D$. Aus der Definition der Varianz und der Linearität des Erwartungswertes folgt für das Semivariogramm

$$\begin{aligned} \gamma(h) &= \frac{1}{2} \text{Var}[Z(u+h) - Z(u)] \\ &= \frac{1}{2} E \left[((Z(u+h) - Z(u)) - E[Z(u+h) - Z(u)])^2 \right] \\ &= \frac{1}{2} E \left[((Z(u+h) - Z(u)) - E[Z(u+h)] + E[Z(u)])^2 \right] \\ &= \frac{1}{2} E \left[((Z(u+h) - Z(u)) - \mu + \mu)^2 \right] \\ &= \frac{1}{2} E \left[(Z(u+h) - Z(u))^2 \right]. \end{aligned} \tag{3.1}$$

Wir beginnen nun das empirische Semivariogramm für die Abstandsvektoren h zu berechnen, die in den Messdaten vorkommen. Gleichung (3.1) folgend wird das empirische Semivariogramm für den Abstandsvektor h definiert als

$$\gamma(h) = \frac{1}{2|N(h)|} \sum_{N(h)} (Z(u_i) - Z(u_j))^2,$$

wobei $N(h) = \{(u_i, u_j) \mid u_i - u_j = h, i, j = 1, \dots, d\}$ und mit $|N(h)|$ die Anzahl der Elemente in $N(h)$ bezeichnet wird. Diese Anzahl ist endlich, da nur endlich viele Stützstellen vorhanden sind. Das empirische Semivariogramm charakterisiert somit den räumlichen Zusammenhang der gemessenen Daten.

Aufgrund von meist irregulären Messnetzen ist die obige Definition in der Praxis untauglich. Es befinden sich oft nur wenige Tupel (u_i, u_j) in der Menge $N(h)$. Einer Faustregel folgend sollten sich jedoch mindestens 30 dieser Tupel in der Menge $N(h)$ zu einem festen Abstandsvektor h befinden (vgl. [19]). Somit schwächt man die Definition etwas ab und fordert

$$\begin{aligned} N(h) &= \{(u_i, u_j) \mid \|h\| - \varepsilon < \|u_i - u_j\| < \|h\| + \varepsilon \text{ und} \\ &\quad \angle(h, e_1) - \delta < \angle(u_j - u_i, e_1) < \angle(h, e_1) + \delta, i, j = 1, \dots, d\}, \end{aligned}$$

wobei $\angle(h, e_1)$ den Winkel zwischen h und dem ersten Einheitsvektor e_1 bezeichnet. Nun wählt man ε und δ so groß, dass obige Faustregel eingehalten wird. Andere Autoren empfehlen wiederum für die Abstandstoleranz den Wert $\varepsilon = \frac{1}{2}\|h\|$ und für die Winkeltoleranz den Wert $\delta = 90^\circ$ (vgl. [28]).

Das so erstellte empirische Semivariogramm liefert allerdings nur Werte für den räumlichen Zusammenhang von Abstandsvektoren h , die in der Messung auftreten. Da wir jedoch eine Darstellung des räumlichen Zusammenhangs für alle Abstandsvektoren h benötigen, folgt nun durch Modellierung des theoretischen Semivariogramms der zweite Schritt der Semivariogrammerstellung.

Die fehlenden Werte des theoretischen Semivariogramms erhält man durch Anpassung einer geeigneten Funktion an die Werte des empirischen Semivariogramms. Wie Satz 3.1 zeigt, muss diese Funktion konditional bedingt negativ semidefinit sein.

Definition 3.4. Sei $D \subseteq \mathbb{R}^n$. Eine Funktion $g : D \times D \rightarrow \mathbb{R}$ heißt konditional bedingt negativ semidefinit, falls

$$\sum_{i=1}^d \sum_{j=1}^d \xi_i g(u_i, u_j) \xi_j \leq 0$$

für alle $d \in \mathbb{N}$, alle $u_1, \dots, u_d \in D$ und alle $\xi_1, \dots, \xi_d \in \mathbb{R}$ mit der zusätzlichen Eigenschaft $\sum_{i=1}^d \xi_i = 0$ gilt.

Satz 3.1. Sei D das Untersuchungsgebiet mit $D \subseteq \mathbb{R}^n$. Jede Semivariogrammfunktion $\gamma : D \times D \rightarrow \mathbb{R}$ ist konditional bedingt negativ semidefinit.

Beweis. Sei also $d \in \mathbb{N}$ beliebig, $u_1, \dots, u_d \in D$ und D das Untersuchungsgebiet. Weiter seien ξ_1, \dots, ξ_d reelle Zahlen so, dass $\sum_{i=1}^d \xi_i = 0$. Für die Varianz $\text{Var}[X]$ einer Zufallsvariable X gilt allgemein

$$\begin{aligned} \text{Var}[X] &= E[(X - E[X])^2] \geq 0 \quad \text{und} \\ \text{Var}[X] &= E[X^2] - (E[X])^2. \end{aligned}$$

Somit folgt mit Hilfe der Eigenschaften des Erwartungswertes

$$\begin{aligned} 0 &\leq \text{Var}\left[\sum_{i=1}^d \xi_i Z(u_i)\right] \\ &= E\left[\left(\sum_{i=1}^d \xi_i Z(u_i)\right)^2\right] - \left(E\left[\sum_{i=1}^d \xi_i Z(u_i)\right]\right)^2 \\ &= E\left[\left(\sum_{i=1}^d \xi_i Z(u_i)\right)^2\right] - \left(\mu \sum_{i=1}^d \xi_i\right)^2 \\ &= E\left[\left(\sum_{i=1}^d \xi_i Z(u_i)\right)^2\right] \\ &= E\left[\sum_{i=1}^d \left(\xi_i Z(u_i) \cdot \sum_{j=1}^d \xi_j Z(u_j)\right)\right] \\ &= E\left[\sum_{i=1}^d \sum_{j=1}^d \xi_i Z(u_i) Z(u_j) \xi_j\right] \\ &= \sum_{i=1}^d \sum_{j=1}^d \xi_i E[Z(u_i) Z(u_j)] \xi_j. \end{aligned}$$

Folglich gilt die Ungleichung

$$\sum_{i=1}^d \sum_{j=1}^d \xi_i E[Z(u_i)Z(u_j)] \xi_j \geq 0,$$

mit deren Hilfe wir die Notwendigkeit der konditional bedingten negativen Semidefinitheit für die Semivariogrammfunktion γ folgern können: Mit Hilfe der Rechenregeln für Varianz und Erwartungswert und aufgrund der Eigenschaft $\sum_{i=1}^d \xi_i = 0$ gilt

$$\begin{aligned} & 2 \cdot \sum_{i=1}^d \sum_{j=1}^d \xi_i \gamma(u_i, u_j) \xi_j \\ &= \sum_{i=1}^d \sum_{j=1}^d \xi_i \text{Var}[Z(u_i) - Z(u_j)] \xi_j \\ &= \sum_{i=1}^d \sum_{j=1}^d \xi_i \left(E[(Z(u_i) - Z(u_j))^2] - (E[Z(u_i) - Z(u_j)])^2 \right) \xi_j \\ &= \sum_{i=1}^d \sum_{j=1}^d \xi_i \left(E[(Z(u_i) - Z(u_j))^2] - (\mu - \mu)^2 \right) \xi_j \\ &= \sum_{i=1}^d \sum_{j=1}^d \xi_i \left(E[Z(u_i)^2] - 2 \cdot E[Z(u_i)Z(u_j)] + E[Z(u_j)^2] \right) \xi_j \\ &= \sum_{i=1}^d \sum_{j=1}^d \xi_i E[Z(u_i)^2] \xi_j - 2 \sum_{i=1}^d \sum_{j=1}^d \xi_i E[Z(u_i)Z(u_j)] \xi_j + \sum_{i=1}^d \sum_{j=1}^d \xi_i E[Z(u_j)^2] \xi_j \\ &\leq \sum_{i=1}^d \sum_{j=1}^d \xi_i E[Z(u_i)^2] \xi_j + \sum_{i=1}^d \sum_{j=1}^d \xi_i E[Z(u_j)^2] \xi_j \\ &= \sum_{i=1}^d \left(\xi_i E[Z(u_i)^2] \sum_{j=1}^d \xi_j \right) + \sum_{j=1}^d \left(\xi_j E[Z(u_j)^2] \sum_{i=1}^d \xi_i \right) = 0. \end{aligned}$$

Insgesamt folgt also

$$\sum_{i=1}^d \sum_{j=1}^d \xi_i \gamma(u_i, u_j) \xi_j \leq 0,$$

d.h. die Notwendigkeit der konditional bedingten negativen Semidefinitheit für die Semivariogrammfunktion.

q.e.d.

Diese Semidefinitheit ist für viele Funktionen oft schwer nachweisbar (vgl. [19]). Aus diesem Grund beschränkt man sich bei der Modellierung des theoretischen Semivariogramms auf eine Familie von Funktionen, für welche dieser Nachweis schon erbracht wurde. Dazu zählen unter anderem das sphärische Modell

$$\gamma_{a,b}(h) = \begin{cases} b \left(1.5 \cdot \frac{h}{a} - 0.5 \cdot \left(\frac{h}{a} \right)^3 \right), & \text{für } h \leq a \\ b, & \text{für } h > a, \end{cases}$$

das exponentielle Modell

$$\gamma_{a,b}(h) = b \left(1 - \exp \left(-\frac{3h}{a} \right) \right)$$

und das Gauß'sche Modell

$$\gamma_{a,b}(h) = b \left(1 - \exp \left(-\frac{3h^2}{a^2} \right) \right).$$

Hierbei ist $a > 0$ jeweils die Aussagenweite und $b > 0$ der Schwellenwert. Abbildung 3.6 zeigt diese drei Funktionen für die Parameter $a = 1$ und $b = 1$ im Vergleich.

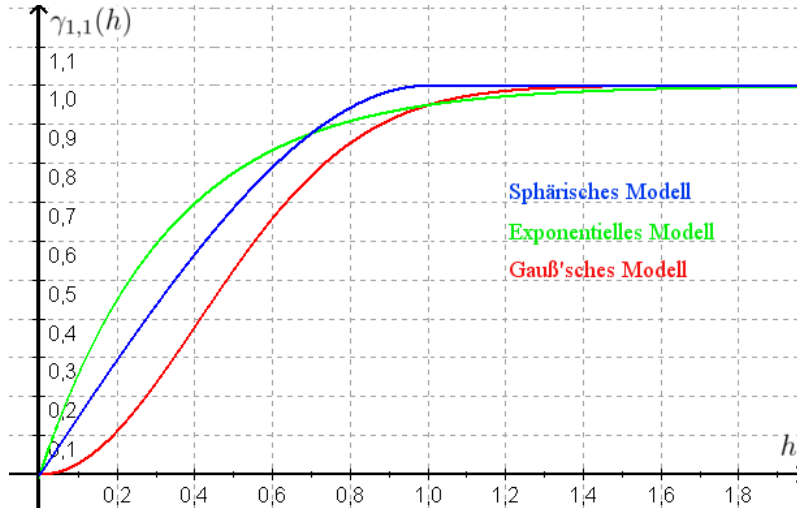


Abbildung 3.6: Verschiedene Semivariogrammfunktionen mit Aussagenweite $a = 1$ und Schwellenwert $b = 1$ im Vergleich.

Die Anpassung einer Funktion an das empirische Semivariogramm erfolgt nun halbautomatisch. Vorerst muss anhand des Verlaufs des empirischen Semivariogramms festgelegt werden, in welcher Familie sich das theoretische Semivariogramm befinden soll. Im Anschluss wird aus dieser Familie das Semivariogramm ausgewählt, welches dem empirischen Semivariogramm am meisten ähnelt. Bei dieser Anpassung der Parameter a und b , dem sogenannten Training[§], kann etwa die Methode der kleinsten Quadrate oder wie im Fall von AutoOpti die Maximum-Likelihood-Methode zum Einsatz kommen.

3.2.2 Ansatz des Kriging-Verfahrens

Im Folgenden sei weiter $D \subseteq \mathbb{R}^n$ das Untersuchungsgebiet, $u_0 \in D$ der unbeprobte Ort und $z(u_0)$ der zu schätzende Wert. Weiter sei die Beobachtungsvariable durch einen stationären stochastischen Prozess $Z = \{Z(u) \mid u \in D\}$ modelliert.

Analog zu nichtstatistischen Verfahren wird bei der Kriging-Methode der zu schätzende Wert $z(u_0)$ an der unbeprobten Stelle u_0 durch ein gewichtetes Mittel der Werte $z(u_1)$ bis $z(u_d)$ an den Stützstellen u_1 bis u_d geschätzt:

$$z^*(u_0) = \sum_{i=1}^d w_i z(u_i).$$

[§]Training in AutoOpti: siehe auch Abb. 2.3, links oben.

Während die Gewichte $w_i \in \mathbb{R}$ bei der Inversen Distanzgewichtung jedoch nur vom Abstand des unbeprobten Ortes u_0 zu den Stützstellen u_1 bis u_d abhängen, basiert die Gewichtung beim Kriging nicht nur auf der Entfernung, sondern auch auf der räumlichen Lage der vermessenen Punkte u_i , $i \in \{1, \dots, d\}$, zueinander (vgl. Abschnitt 3.2.3).

Definition 3.5. Seien $w_i \in \mathbb{R}$ die noch zu bestimmenden Kriging-Gewichte. Der Krige-Schätzer $Z^*(u_0)$ ist definiert als

$$Z^*(u_0) = \sum_{i=1}^d w_i Z(u_i).$$

Weiter sei der Schätzfehler $F(u_0)$ definiert als

$$F(u_0) = Z(u_0) - Z^*(u_0) = Z(u_0) - \sum_{i=1}^d w_i Z(u_i).$$

Beim Krige-Schätzer handelt es sich um einen BLUE-Schätzer (vgl. S. 15). Somit sind die Kriging-Gewichte w_i so zu wählen, dass der Schätzer im Mittel den wahren Wert schätzt und keinen systematischen Fehler macht. Folglich ist der Schätzfehler im Mittel gleich Null, d.h.

$$E[F(u_0)] = 0.$$

Daraus folgt

$$E[Z^*(u_0)] = E[Z(u_0)].$$

Der Schätzer ist somit erwartungstreu. Weiter soll die Varianz des Schätzfehlers, die sogenannte Kriging-Varianz, minimal sein, d.h.

$$\text{Var}[F(u_0)] = \min_{w_1, \dots, w_d} \text{Var}\left[Z(u_0) - \sum_{i=1}^d w_i Z(u_i)\right].$$

Zur Bestimmung der Kriging-Gewichte gilt es also, die Extremwertaufgabe

$$(E) \begin{cases} \min_{w_1, \dots, w_d} \text{Var}\left[Z(u_0) - \sum_{i=1}^d w_i Z(u_i)\right], \\ \text{sodass } E[F(u_0)] = 0 \end{cases}$$

zu lösen.

3.2.3 Ordinary Kriging

Im Fall des Ordinary Kriging können wir wegen der geforderten Bedingung $E[Z(u)] = \mu$ für alle $u \in D$ die Nebenbedingung $E[F(u_0)] = 0$ zu $\sum_{i=1}^d w_i = 1$ vereinfachen, denn es gilt

$$\begin{aligned}
0 &= E[F(u_0)] \\
&= E\left[Z(u_0) - \sum_{i=1}^d w_i Z(u_i)\right] \\
&= E[Z(u_0)] - \sum_{i=1}^d w_i E[Z(u_i)] \\
&= \mu \left(1 - \sum_{i=1}^d w_i\right).
\end{aligned}$$

Da $\mu \in \mathbb{R}$ fest, aber je nach Art des stationären stochastischen Prozesses verschieden ist, folgt somit tatsächlich

$$\sum_{i=1}^d w_i = 1.$$

Mit Hilfe dieser Nebenbedingung können wir nun auch die zu minimierende Varianz des Schätzfehlers vereinfachen. Wir zeigen

$$\text{Var}[F(u_0)] = 2 \sum_{i=1}^d w_i \gamma(u_i, u_0) - \sum_{i=1}^d \sum_{j=1}^d w_i \gamma(u_i, u_j) w_j. \quad (3.2)$$

Beweis. Aufgrund der Nebenbedingung $\sum_{i=1}^d w_i = 1$ und mit Hilfe der schon mehrfach angewandten Rechenregeln für Erwartungswert und Varianz folgt

$$\begin{aligned}
&\text{Var}[F(u_0)] \\
&= \text{Var}\left[Z(u_0) - \sum_{i=1}^d w_i Z(u_i)\right] \\
&= E\left[\left(Z(u_0) - \sum_{i=1}^d w_i Z(u_i)\right)^2\right] - \left(E\left[Z(u_0) - \sum_{i=1}^d w_i Z(u_i)\right]\right)^2 \\
&= E\left[\left(Z(u_0) - \sum_{i=1}^d w_i Z(u_i)\right)^2\right] - \mu \left(1 - \sum_{i=1}^d w_i\right) \\
&= E[Z(u_0)^2] - 2E\left[Z(u_0) \cdot \left(\sum_{i=1}^d w_i Z(u_i)\right)\right] + E\left[\left(\sum_{i=1}^d w_i Z(u_i)\right)^2\right] \\
&= E[Z(u_0)^2] - 2 \sum_{i=1}^d w_i E[Z(u_0) Z(u_i)] + \sum_{i=1}^d \sum_{j=1}^d w_i E[Z(u_i) Z(u_j)] w_j. \quad (3.3)
\end{aligned}$$

Betrachte nun die zu beweisende rechte Seite:

$$\begin{aligned}
& 2 \sum_{i=1}^d w_i \gamma(u_i, u_0) - \sum_{i=1}^d \sum_{j=1}^d w_i \gamma(u_i, u_j) w_j \\
&= \sum_{i=1}^d w_i \text{Var}[Z(u_i) - Z(u_0)] - \frac{1}{2} \sum_{i=1}^d \sum_{j=1}^d w_i \text{Var}[Z(u_i) - Z(u_j)] w_j \\
&= \sum_{i=1}^d w_i \left(E[(Z(u_i) - Z(u_0))^2] - (E[Z(u_i) - Z(u_0)])^2 \right) \\
&\quad - \frac{1}{2} \sum_{i=1}^d \sum_{j=1}^d w_i \left(E[(Z(u_i) - Z(u_j))^2] - (E[Z(u_i) - Z(u_j)])^2 \right) w_j \\
&= \sum_{i=1}^d w_i \left(E[(Z(u_i) - Z(u_0))^2] - (\mu - \mu)^2 \right) \\
&\quad - \frac{1}{2} \sum_{i=1}^d \sum_{j=1}^d w_i \left(E[(Z(u_i) - Z(u_j))^2] - (\mu - \mu)^2 \right) w_j \\
&= \sum_{i=1}^d w_i \left(E[Z(u_i)^2] - 2E[Z(u_0)Z(u_i)] + E[Z(u_0)^2] \right) \\
&\quad - \frac{1}{2} \sum_{i=1}^d \sum_{j=1}^d w_i \left(E[Z(u_i)^2] - 2E[Z(u_i)Z(u_j)] + E[Z(u_j)^2] \right) w_j \\
&= \sum_{i=1}^d w_i E[Z(u_i)^2] - 2 \sum_{i=1}^d w_i E[Z(u_0)Z(u_i)] + E[Z(u_0)^2] \left(\sum_{i=1}^d w_i \right) \\
&\quad - \frac{1}{2} \sum_{i=1}^d \left(w_i E[Z(u_i)^2] \sum_{j=1}^d w_j \right) + \sum_{i=1}^d \sum_{j=1}^d w_i E[Z(u_i)Z(u_j)] w_j \\
&\quad - \frac{1}{2} \sum_{j=1}^d \left(w_j E[Z(u_j)^2] \sum_{i=1}^d w_i \right) \\
&= \sum_{i=1}^d w_i E[Z(u_i)^2] - 2 \sum_{i=1}^d w_i E[Z(u_0)Z(u_i)] + E[Z(u_0)^2] \\
&\quad - \sum_{i=1}^d w_i E[Z(u_i)^2] + \sum_{i=1}^d \sum_{j=1}^d w_i E[Z(u_i)Z(u_j)] w_j \\
&= E[Z(u_0)^2] - 2 \sum_{i=1}^d w_i E[Z(u_0)Z(u_i)] + \sum_{i=1}^d \sum_{j=1}^d w_i E[Z(u_i)Z(u_j)] w_j.
\end{aligned}$$

Mit (3.3) folgt somit die zu beweisende Gleichung

$$\text{Var}[F(u_0)] = 2 \sum_{i=1}^d w_i \gamma(u_i, u_0) - \sum_{i=1}^d \sum_{j=1}^d w_i \gamma(u_i, u_j) w_j.$$

q.e.d.

Das obige Extremwertproblem (E) lässt sich folglich zu

$$(E) \left\{ \begin{array}{l} \min_{w_1, \dots, w_d} f(w_1, \dots, w_d) := 2 \sum_{i=1}^d w_i \gamma(u_i, u_0) - \sum_{i=1}^d \sum_{j=1}^d w_i \gamma(u_i, u_j) w_j, \\ \text{sodass } \sum_{i=1}^d w_i - 1 = 0 \end{array} \right.$$

umformulieren. Diese Extremwertaufgabe mit Nebenbedingung können wir mit Hilfe der Lagrange'schen Multiplikatorenregel lösen. Sei im Folgenden $\lambda \in \mathbb{R}$ der Lagrange-Multiplikator. Für die partielle Ableitung von f gilt

$$\frac{\partial f}{\partial w_j}(w_1, \dots, w_d) = 2\gamma(u_j, u_0) - 2 \sum_{i=1}^d w_i \gamma(u_j, u_i)$$

für $j = 1, \dots, d$. Wir erhalten sodann das Gleichungssystem

$$\left\{ \begin{array}{l} 2\gamma(u_j, u_0) - 2 \sum_{i=1}^d w_i \gamma(u_j, u_i) = \lambda \quad \forall j = 1, \dots, d \\ \sum_{i=1}^d w_i - 1 = 0, \end{array} \right.$$

welches wir mit $\nu := \frac{\lambda}{2}$ äquivalent zu

$$\left\{ \begin{array}{l} \sum_{i=1}^d w_i \gamma(u_j, u_i) + \nu = \gamma(u_j, u_0) \quad \forall j = 1, \dots, d \\ \sum_{i=1}^d w_i = 1, \end{array} \right. \quad (3.4)$$

$$\sum_{i=1}^d w_i = 1, \quad (3.5)$$

umformulieren können. In Matrix-Schreibweise erhalten wir

$$\begin{pmatrix} \gamma(u_1, u_1) & \dots & \gamma(u_1, u_d) & 1 \\ \vdots & \ddots & \vdots & \vdots \\ \gamma(u_d, u_1) & \dots & \gamma(u_d, u_d) & 1 \\ 1 & \dots & 1 & 0 \end{pmatrix} \begin{pmatrix} w_1 \\ \vdots \\ w_d \\ \nu \end{pmatrix} = \begin{pmatrix} \gamma(u_1, u_0) \\ \vdots \\ \gamma(u_d, u_0) \\ 1 \end{pmatrix}. \quad (3.6)$$

Im weiteren Verlauf des Abschnitts verwenden wir die Notation $\Gamma = (\gamma(u_i, u_j))_{i,j=1,\dots,d}$, $\mathbf{w} = (w_1, \dots, w_d)^T$ und $\boldsymbol{\gamma} = (\gamma(u_1, u_0), \dots, \gamma(u_d, u_0))^T$. Weiter sei der d -dimensionale Einervektor mit $\mathbf{1}$ bezeichnet.

Wegen Satz 3.1 ist die obige Matrix invertierbar (vgl. [29]). Wir können also

$$\begin{pmatrix} \mathbf{w} \\ \nu \end{pmatrix} = \begin{pmatrix} \Gamma & \mathbf{1} \\ \mathbf{1}^T & 0 \end{pmatrix}^{-1} \begin{pmatrix} \boldsymbol{\gamma} \\ 1 \end{pmatrix} \quad (3.7)$$

schreiben und die Kriging-Gewichte für $Z^*(u_0)$ sind eindeutig bestimmt. Wir erhalten somit den Krige-Schätzer am unbeprobten Ort u_0 durch Berechnung von

$$Z^*(u_0) = \sum_{i=1}^d w_i Z(u_i).$$

In Anlehnung an die in der Literatur verwendete Schreibweise für den Krige-Schätzer (vgl. [11]) werden wir im Folgenden mit Hilfe von Formel (3.7) weitere Umformulierungen durchführen, sodass der Schätzer letztendlich explizit angegeben werden kann. Nach Jang [30] gilt

$$\begin{pmatrix} \Gamma & \mathbf{1} \\ \mathbf{1}^T & 0 \end{pmatrix}^{-1} = \begin{pmatrix} \Gamma^{-1} + \frac{1}{k} \Gamma^{-1} \mathbf{1} \mathbf{1}^T \Gamma^{-1} & -\frac{1}{k} \Gamma^{-1} \mathbf{1} \\ -\frac{1}{k} \mathbf{1}^T \Gamma^{-1} & \frac{1}{k} \end{pmatrix}$$

mit $k = -\mathbf{1}^T \Gamma^{-1} \mathbf{1}$. Somit folgt für die Kriging-Gewichte \mathbf{w} in Gleichung (3.7)

$$\mathbf{w} = \left(\Gamma^{-1} + \frac{1}{k} \Gamma^{-1} \mathbf{1} \mathbf{1}^T \Gamma^{-1} \right) \boldsymbol{\gamma} - \frac{1}{k} \Gamma^{-1} \mathbf{1}.$$

Sei im Folgenden $\mathbf{Z} = (Z(u_1), \dots, Z(u_d))^T$, so erhalten wir für den Krige-Schätzer

$$\begin{aligned} Z^*(u_0) &= \sum_{i=1}^d w_i Z(u_i) \\ &= \mathbf{w}^T \mathbf{Z} \\ &= \left[\left(\Gamma^{-1} + \frac{1}{k} \Gamma^{-1} \mathbf{1} \mathbf{1}^T \Gamma^{-1} \right) \boldsymbol{\gamma} - \frac{1}{k} \Gamma^{-1} \mathbf{1} \right]^T \mathbf{Z}. \end{aligned}$$

Da die Matrix Γ symmetrisch ist und allgemein $(\Gamma^{-1})^T = (\Gamma^T)^{-1}$ gilt, können wir weiter folgern, dass

$$\begin{aligned} Z^*(u_0) &= \left(\boldsymbol{\gamma}^T \Gamma^{-1} + \frac{1}{k} \boldsymbol{\gamma}^T \Gamma^{-1} \mathbf{1} \mathbf{1}^T \Gamma^{-1} - \frac{1}{k} \mathbf{1}^T \Gamma^{-1} \right) \mathbf{Z} \\ &= \boldsymbol{\gamma}^T \Gamma^{-1} \left(\mathbf{Z} + \frac{1}{k} \mathbf{1} \mathbf{1}^T \Gamma^{-1} \mathbf{Z} \right) - \frac{1}{k} \mathbf{1}^T \Gamma^{-1} \mathbf{Z} \\ &= \boldsymbol{\gamma}^T \Gamma^{-1} (\mathbf{Z} - \mathbf{1} \cdot \beta) + \beta \\ &= \beta + \boldsymbol{\gamma}^T \Gamma^{-1} (\mathbf{Z} - \beta \cdot \mathbf{1}), \end{aligned} \tag{3.8}$$

mit $\beta = -\frac{1}{k} \mathbf{1}^T \Gamma^{-1} \mathbf{Z} = \frac{\mathbf{1}^T \Gamma^{-1} \mathbf{Z}}{\mathbf{1}^T \Gamma^{-1} \mathbf{1}}$. Da für jeden unbeprobten Ort die Kriging-Gewichte neu berechnet werden müssen, wird die Invertierung der Matrix Γ tatsächlich auch in der Praxis vorgenommen.

Betrachten wir das obige Gleichungssystem (3.6) respektive (3.7) genauer, so erkennen wir einen der Vorteile des Kriging gegenüber einem Verfahren wie der Inversen Distanzgewichtung. Der Vektor $\boldsymbol{\gamma}$ enthält die Werte des räumlichen Zusammenhangs zwischen dem unbeprobten Ort u_0 und den Stützstellen u_i , $i = 1, \dots, d$. Würden die Kriging-Gewichte nur mit Hilfe dieses Vektors bestimmt werden, so entstünde ein der Inversen Distanzgewichtung ähnliches Verfahren. Der geometrische Abstand wird allerdings durch das Semivariogramm ersetzt (vgl. [26]).

Die Matrix Γ liefert nun zusätzlich die Werte des räumlichen Zusammenhangs der Stützstellen untereinander. Durch die Multiplikation von Γ^{-1} mit dem Vektor $\boldsymbol{\gamma}$ werden Cluster, also nah beieinander liegende Stützstellen erfasst und die Gewichte dieser gehäuften Stützstellen entsprechend reduziert. Insbesondere können auch negative Gewichte auftreten. Dies geschieht beispielsweise bei starker Clusterung oder wenn die Messdaten bei kleinen Distanzen einen hohen Zusammenhang aufweisen, wie es etwa beim Gauß'schen Modell der Fall ist (vgl. [19]).

Für die Varianz des Schätzfehlers, oft auch als Kriging- oder Fehlervarianz bezeichnet, erhalten wir daraufhin im Minimum

$$\text{Var}[F(u_0)] = \sum_{i=1}^d w_i \gamma(u_i, u_0) + \nu = \mathbf{w}^T \boldsymbol{\gamma} + \nu, \quad (3.9)$$

denn aus Gleichung (3.4) folgt durch Multiplikation mit w_j und anschließender Addition über alle $j = 1, \dots, d$

$$\sum_{j=1}^d \sum_{i=1}^d w_i \gamma(u_j, u_i) w_j + \nu \sum_{j=1}^d w_j = \sum_{j=1}^d \gamma(u_j, u_0) w_j.$$

Aufgrund von (3.5) folgt sodann

$$\sum_{j=1}^d \sum_{i=1}^d w_i \gamma(u_j, u_i) w_j = \sum_{j=1}^d \gamma(u_j, u_0) w_j - \nu.$$

Einsetzen in Formel (3.2) liefert das obige Resultat (3.9) für die Varianz des Schätzfehlers.

3.2.4 Weitere Kriging-Methoden

Wie bereits erwähnt existieren verschiedenste Kriging-Methoden. Abhängig von den stochastischen Eigenschaften der Zufallsvariablen und des zugrundeliegenden stochastischen Prozesses wird die jeweils passende Kriging-Methode ausgewählt.

So war beispielsweise beim Ordinary Kriging die Stationarität des stochastischen Prozesses gefordert. Unter anderem galt für den Erwartungswert, dass dieser konstant für alle Zufallsvariablen ist, d.h. $\exists \mu \in \mathbb{R}$, sodass

$$E[Z(u)] = \mu \quad \forall u \in D.$$

Im Fall des Ordinary Kriging handelt es sich um einen nicht bekannten, konstanten Erwartungswert μ .

Eine Vereinfachung des Ordinary Kriging bildet das Simple Kriging. Auch hier wird die Stationarität des stochastischen Prozesses gefordert. Allerdings gehen wir nun von einem bekannten, konstanten Erwartungswert $\mu \equiv 0$ aus. Um das Simple Kriging anwenden zu können, müssen also mehr Informationen über die Zufallsvariable, nämlich gerade der Erwartungswert dieser, bekannt sein.

Beim Simple Kriging handelt es sich um die einfachste Form des Kriging, jedoch um die am wenigsten allgemeine. Durch die Forderung des bekannten Erwartungswertes ist eine starke Einschränkung gegeben, da dieser in den meisten Fällen unbekannt ist (vgl. [25]).

Das Universal Kriging ist eine Erweiterung des Ordinary Kriging. Der Erwartungswert der Zufallsvariablen ist nun nicht mehr konstant; es ist ein Trend in den Messdaten vorhanden. Dieser kann sich über das gesamte Untersuchungsgebiet D erstrecken oder nur lokal vorhanden sein. Ersteres bezeichnet man als globalen Trend, letzteres als sogenannte lokale Drift.

Wir erhalten somit für den Erwartungswert die Darstellung

$$E[Z(u)] = \mu(u).$$

Die zu bestimmende Funktion $\mu : D \rightarrow \mathbb{R}$ ist dabei eine deterministische Funktion, welche sich als Linearkombination von „einfachen“ Funktionen $f_k : D \rightarrow \mathbb{R}$ schreiben lässt:

$$\mu(u) = \sum_{k=0}^p \beta_k f_k(u).$$

Hierbei ist $\beta_k \in \mathbb{R}$ für $k = 0, \dots, p$ und $p \in \mathbb{N}$. Die Funktionen f_k können beispielsweise Polynome sein (vgl. [27, 31]).

Neben den drei oben genannten Verfahren existieren noch zahlreiche weitere Kriging-Methoden. Auszugsweise seien hier das IRFk-Kriging, das Indicator Kriging und das Disjunctive Kriging genannt. Auf diese Methoden wollen wir im Rahmen dieser Arbeit nicht weiter eingehen. Es sei auf Chilès und Delfiner [29] verwiesen.

3.2.5 Gradient Enhanced Kriging

Bei der Gradient Enhanced Kriging-Methode (GEK) handelt es sich um eine erweiterte Form des Ordinary Kriging. Hierbei fließen vorhandene Gradienteninformationen der zu interpolierenden Funktion an einigen der beprobten Orte in die Modellbildung ein.

Im Fall von AutoOpti erhalten wir diese Gradienteninformationen beziehungsweise Formableitungen mit Hilfe des Tools adjointTRACE, welches die diskret adjungierten Reynolds-Gleichungen löst (vgl. Kapitel 1).

Es existieren zwei verschiedene Möglichkeiten, diese Gradienteninformationen bei der Modellbildung einzubeziehen: das Direct GEK und das Indirect GEK (vgl. [32]). Im Folgenden werden wir beide Methoden kurz vorstellen. Insbesondere das Direct GEK soll genauer untersucht werden, da dieses Verfahren zur Erstellung der Ersatzmodelle in AutoOpti Anwendung findet. Wir orientieren uns an Yamazaki et al. [32] und Backhaus et al. [11].

Im Folgenden sei weiterhin d die Anzahl der Stützstellen und $D \subseteq \mathbb{R}^n$ das Untersuchungsgebiet. Weiter seien die partiellen Ableitungen an $\tilde{d} \leq d$ Stützstellen gegeben. Dabei handelt es sich o.B.d.A. um die Stützstellen u_1 bis $u_{\tilde{d}}$.

Die Idee des Indirect GEK besteht darin, mit Hilfe der Gradienteninformationen weitere Stützstellen in der Umgebung der vorhandenen und mit Gradienteninformationen ausgestatteten Stützstellen u_1 bis $u_{\tilde{d}}$ zu erzeugen.

Sei u_i , $i \in \{1, \dots, \tilde{d}\}$, die ursprüngliche Stützstelle und $u_{i_{\text{add}}}$ die neue Stützstelle nahe u_i so, dass $u_{i_{\text{add}}} = u_i + \Delta u$. Die Ermittlung des Funktionswertes $z(u_{i_{\text{add}}})$ an dieser neuen Stützstelle erfolgt nun mit Hilfe einer Taylor-Approximation erster Ordnung

$$z(u_{i_{\text{add}}}) = z(u_i) + \Delta u^T \nabla z(u_i).$$

Üblicherweise wird mit dieser Technik je ein neuer Punkt in jeder Raumrichtung erzeugt. Wir erhalten somit pro Stützstelle n neue Stützstellen, insgesamt erhöht sich die Anzahl der Stützstellen um $\tilde{d} \cdot n$. Im Anschluss an die Stützstellenerzeugung kann man nun wie gewohnt mit dem Ordinary Kriging fortfahren.

Ein Vorteil des Indirect GEK besteht in der relativ einfachen Implementierung. Einen entscheidenden Nachteil stellen die zusätzlichen Stützstellen jedoch dar: Vor allem für kleine Schrittweiten Δu kann das Indirect GEK aufgrund der gehäuften und nah beieinanderliegenden Stützstellen eine schlecht konditionierten Matrix ergeben.

Im Gegensatz zum Indirect GEK fließen die Gradienteninformationen beim Direct GEK unmittelbar in die Modellbildung ein. Es werden dabei nicht nur die räumlichen Zusammenhänge $Var[Z(x) - Z(y)]$ zwischen zwei Zufallsvariablen betrachtet, sondern auch die Zusammenhänge zwischen einer Zufallsvariable und der partiellen Ableitung einer anderen (vgl. Formel (3.11)) sowie zwischen beiden partiellen Ableitungen (vgl. Formel (3.12)). Diese zusätzlichen räumlichen Zusammenhänge erhalten wir durch entsprechendes Ableiten der Semivariogrammfunktion.

Die nachfolgenden Formeln und deren Beweis konnten nicht in der Literatur gefunden werden und beruhen daher auf eigenen Überlegungen.

Seien $k, m \in \{1, \dots, n\}$ und $x, y \in D \subseteq \mathbb{R}^n$. Aus Gründen der Übersichtlichkeit verwenden wir im Folgenden die Notation $Z'(x) = \frac{\partial}{\partial x^m} Z(x)$ und $Z'(y) = \frac{\partial}{\partial y^k} Z(y)$ für die partiellen Ableitungen von Z .

Für zwei Zufallsvariablen X und Y gilt die Formel

$$Var[X - Y] = Var[X] + Var[Y] - 2 \cdot Cov(X, Y), \quad (3.10)$$

welche aus der Definition der Kovarianz und mit Hilfe der Rechenregeln für Erwartungswert und Varianz hergeleitet werden kann:

$$\begin{aligned} Var[X - Y] &= Var[X] + Var[-Y] + 2 \cdot Cov(X, -Y) \\ &= Var[X] + E[(-Y + E[-Y])^2] + 2 \cdot E[(X - E[X])(-Y - E[-Y])] \\ &= Var[X] + E[(Y + E[Y])^2] - 2 \cdot E[(X - E[X])(Y - E[Y])] \\ &= Var[X] + Var[Y] - 2 \cdot Cov(X, Y). \end{aligned}$$

Somit gilt

$$\begin{aligned} Var[Z'(x) - Z'(y)] &= Var[Z'(x)] + Var[Z'(y)] - 2 \cdot Cov(Z'(x), Z'(y)) \\ &= Cov(Z'(x), Z'(x)) + Cov(Z'(y), Z'(y)) - 2 \cdot Cov(Z'(x), Z'(y)) \end{aligned} \quad (3.11)$$

und

$$\begin{aligned} Var[Z'(x) - Z'(y)] &= Var[Z'(x)] + Var[Z'(y)] - 2 \cdot Cov(Z'(x), Z'(y)) \\ &= Cov(Z'(x), Z'(x)) + Cov(Z'(y), Z'(y)) - 2 \cdot Cov(Z'(x), Z'(y)). \end{aligned} \quad (3.12)$$

Die jeweils zweite Gleichheit folgt aufgrund der Definition der Kovarianz. Diese Kovarianzen können wir als Ableitungen der uns bekannten Semivariogrammfunktion γ darstellen, denn es gilt

$$\begin{aligned} Cov(Z'(x), Z'(y)) &= -\frac{1}{2} \cdot \frac{\partial^2}{\partial x^m \partial y^k} Var[Z(x) - Z(y)] \\ &= -\frac{\partial^2}{\partial x^m \partial y^k} \gamma(x, y) \end{aligned} \quad (3.13)$$

und

$$\begin{aligned} Cov(Z'(x), Z(y)) &= -\frac{1}{2} \cdot \frac{\partial}{\partial x^m} Var[Z(x) - Z(y)] \\ &= -\frac{\partial}{\partial x^m} \gamma(x, y). \end{aligned} \quad (3.14)$$

Beweis. Mit Hilfe der zuvor bewiesenen Formel (3.10) und aufgrund der Unabhängigkeit der beiden Variablen folgt

$$\begin{aligned}
\frac{\partial^2}{\partial x^m \partial y^k} \gamma(x, y) &= \frac{\partial^2}{\partial x^m \partial y^k} \frac{1}{2} \cdot \text{Var}[Z(x) - Z(y)] \\
&= \frac{\partial^2}{\partial x^m \partial y^k} \frac{1}{2} \cdot \left(\text{Var}[Z(x)] + \text{Var}[Z(y)] - 2 \cdot \text{Cov}(Z(x), Z(y)) \right) \\
&= - \frac{\partial^2}{\partial x^m \partial y^k} \text{Cov}(Z(x), Z(y)) \\
&= - \frac{\partial^2}{\partial x^m \partial y^k} E \left[(Z(x) - E[Z(x)])(Z(y) - E[Z(y)]) \right] \\
&= - E \left[\frac{\partial^2}{\partial x^m \partial y^k} \left((Z(x) - E[Z(x)])(Z(y) - E[Z(y)]) \right) \right] \\
&= - E \left[\frac{\partial}{\partial x^m} (Z(x) - E[Z(x)]) \frac{\partial}{\partial y^k} (Z(y) - E[Z(y)]) \right] \\
&= - E \left[(Z'(x) - E[Z'(x)])(Z'(y) - E[Z'(y)]) \right] \\
&= - \text{Cov}(Z'(x), Z'(y)).
\end{aligned}$$

Somit folgt die erste zu beweisende Gleichung (3.13). Weiter gilt

$$\begin{aligned}
\frac{\partial}{\partial x^m} \gamma(x, y) &= \frac{\partial}{\partial x^m} \frac{1}{2} \cdot \text{Var}[Z(x) - Z(y)] \\
&= \frac{\partial}{\partial x^m} \frac{1}{2} \cdot \left(\text{Var}[Z(x)] + \text{Var}[Z(y)] - 2 \cdot \text{Cov}(Z(x), Z(y)) \right) \\
&= \frac{\partial}{\partial x^m} \frac{1}{2} \cdot \text{Var}[Z(x)] - \frac{\partial}{\partial x^m} \text{Cov}(Z(x), Z(y)) \\
&= - \frac{\partial}{\partial x^m} \text{Cov}(Z(x), Z(y)) \\
&= - \frac{\partial}{\partial x^m} E \left[(Z(x) - E[Z(x)])(Z(y) - E[Z(y)]) \right] \\
&= - E \left[\frac{\partial}{\partial x^m} \left((Z(x) - E[Z(x)])(Z(y) - E[Z(y)]) \right) \right] \\
&= - E \left[\left(\frac{\partial}{\partial x^m} (Z(x) - E[Z(x)]) \right) (Z(y) - E[Z(y)]) \right] \\
&= - E \left[(Z'(x) - E[Z'(x)])(Z(y) - E[Z(y)]) \right] \\
&= - \text{Cov}(Z'(x), Z(y)),
\end{aligned}$$

wobei die vierte Gleichheit aufgrund der Stationarität des stochastischen Prozesses Z und der daraus folgenden Konstanz der Varianz $\text{Var}[Z(u)]$ gefolgert werden kann.

Die mehrfach angewandte Vertauschbarkeit von Differentiation und Erwartungswert folgt aufgrund der Vertauschbarkeit von Differentiation und Integration, welche unter Erfüllung der notwendigen Voraussetzungen wiederum aus Lebegues Satz der dominierten Konvergenz gefolgert werden kann. Die Voraussetzungen zur Anwendung von Lebesgues Konvergenzsatz sind hier erfüllt.

q.e.d.

Mit Hilfe der Formeln (3.11) und (3.12) sowie (3.13) und (3.14) können wir somit die gesuchten räumlichen Zusammenhänge als Ableitung der Semivariogrammfunktion γ berechnen.

Der Ansatz für den Krige-Schätzer lautet nun (vgl. [32])

$$Z^*(u_0) = \sum_{i=1}^d w_i Z(u_i) + \sum_{i=1}^{\tilde{d}} \sum_{j=1}^n \Phi_{ij} \frac{\partial Z(u_i)}{\partial u^j},$$

wobei $\frac{\partial Z(u_i)}{\partial u^j}$ die gegebenen partiellen Ableitungen an den Stützstellen u_1 bis $u_{\tilde{d}}$ mit den Gewichten Φ_{ij} , $i \in \{1, \dots, \tilde{d}\}$, $j \in \{1, \dots, n\}$, sind.

Die geforderte Minimierung der Varianz des Schätzfehlers

$$F(u_0) = Z(u_0) - Z^*(u_0) = Z(u_0) - \sum_{i=1}^d w_i Z(u_i) - \sum_{i=1}^{\tilde{d}} \sum_{j=1}^n \Phi_{ij} \frac{\partial Z(u_i)}{\partial u^j}$$

unter der Nebenbedingung der Erwartungstreue

$$E[Z^*(u_0)] = E[Z(u_0)]$$

führt erneut zu einem Extremwertproblem mit Nebenbedingung, welches wir wiederum mit Hilfe der Lagrange'schen Multiplikatorenregel lösen.

Sei im Folgenden ν der Lagrange-Multiplikator. Wir erhalten schließlich das folgende um V_2 und V_3 erweiterte Gleichungssystem

$$\begin{pmatrix} & & & 1 \\ & V_1 & & V_2 & \vdots \\ & & & 1 \\ & & & 0 \\ V_2^T & & V_3 & \vdots \\ 1 & \dots & 1 & 0 & \dots & 0 \end{pmatrix} \begin{pmatrix} w_1 \\ \vdots \\ w_d \\ \Phi_{11} \\ \vdots \\ \Phi_{\tilde{d}n} \\ \nu \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \text{Var}[Z(u_1) - Z(u_0)] \\ \vdots \\ \frac{1}{2} \text{Var}[Z(u_d) - Z(u_0)] \\ \frac{1}{2} \text{Var}\left[\frac{\partial Z(u_1)}{\partial u^1} - Z(u_0)\right] \\ \vdots \\ \frac{1}{2} \text{Var}\left[\frac{\partial Z(u_{\tilde{d}})}{\partial u^n} - Z(u_0)\right] \\ 1 \end{pmatrix},$$

wobei $V_1 \in \mathbb{R}^{d \times d}$, $V_2 \in \mathbb{R}^{d \times \tilde{d} \cdot n}$ und $V_3 \in \mathbb{R}^{\tilde{d} \cdot n \times \tilde{d} \cdot n}$. Genauer ist

$$V_1 = \begin{pmatrix} \frac{1}{2} \text{Var}[Z(u_1) - Z(u_1)] & \dots & \frac{1}{2} \text{Var}[Z(u_1) - Z(u_d)] \\ \vdots & \ddots & \vdots \\ \frac{1}{2} \text{Var}[Z(u_d) - Z(u_1)] & \dots & \frac{1}{2} \text{Var}[Z(u_d) - Z(u_d)] \end{pmatrix}$$

und

$$V_2 = (V_{2_1} \quad \dots \quad V_{2_{\tilde{d}}}),$$

mit

$$V_{2_i} = \begin{pmatrix} \frac{1}{2} \text{Var}\left[Z(u_1) - \frac{\partial Z(u_i)}{\partial u^1}\right] & \dots & \frac{1}{2} \text{Var}\left[Z(u_1) - \frac{\partial Z(u_i)}{\partial u^n}\right] \\ \vdots & \ddots & \vdots \\ \frac{1}{2} \text{Var}\left[Z(u_d) - \frac{\partial Z(u_i)}{\partial u^1}\right] & \dots & \frac{1}{2} \text{Var}\left[Z(u_d) - \frac{\partial Z(u_i)}{\partial u^n}\right] \end{pmatrix}, \quad i \in \{1, \dots, \tilde{d}\}.$$

Die Matrix V_3 besteht entsprechend aus den Einträgen

$$\frac{1}{2} \text{Var}\left[\frac{\partial Z(u_i)}{\partial u^m} - \frac{\partial Z(u_j)}{\partial u^k}\right],$$

wobei $i, j \in \{1, \dots, \tilde{d}\}$ und $m, k \in \{1, \dots, n\}$.

Durch Invertieren werden wiederum die Gewichte w_1 bis w_d sowie Φ_{11} bis $\Phi_{\tilde{d}n}$ bestimmt. Wir erhalten somit für den Krige-Schätzer am unbeprobten Ort u_0 die Darstellung

$$Z^*(u_0) = \sum_{i=1}^d w_i Z(u_i) + \sum_{i=1}^{\tilde{d}} \sum_{j=1}^n \Phi_{ij} \frac{\partial Z(u_i)}{\partial u^j}$$

mit bekannten Gewichten w_i und Φ_{ij} . Bezeichnen wir die erweiterte Semivariogramm-Matrix

$$\begin{pmatrix} V_1 & V_2 \\ V_2^T & V_3 \end{pmatrix}$$

mit Γ und verwenden die folgenden Bezeichnungen

$$\begin{aligned} \mathbf{Z} &= \left(Z(u_1), \dots, Z(u_d), \frac{\partial Z(u_1)}{\partial u^1}, \dots, \frac{\partial Z(u_{\tilde{d}})}{\partial u^n} \right)^T, \\ \mathbf{1} &= (1, \dots, 1, 0, \dots, 0)^T \text{ und} \\ \boldsymbol{\gamma} &= \left(\frac{1}{2} \text{Var}[Z(u_1) - Z(u_0)], \dots, \frac{1}{2} \text{Var}[Z(u_d) - Z(u_0)], \right. \\ &\quad \left. \frac{1}{2} \text{Var}\left[\frac{\partial Z(u_1)}{\partial u^1} - Z(u_0)\right], \dots, \frac{1}{2} \text{Var}\left[\frac{\partial Z(u_{\tilde{d}})}{\partial u^n} - Z(u_0)\right] \right)^T \end{aligned}$$

für die erweiterten Vektoren, so erhalten wir für den Krige-Schätzer im Direct Gradient Enhanced Kriging dieselbe Darstellung wie in (3.8) auf Seite 27 (vgl. [11]).

4 Nichtlineare Optimierung unter Nebenbedingungen

Ziel dieser Arbeit ist die Optimierung von Ersatzmodellen unter Beachtung eventuell vorhandener Nebenbedingungen. Diese werden an die mit Hilfe des Direct GEK berechneten Unsicherheiten oder andere aerodynamische Größen gestellt (vgl. Kapitel 5). Aufgrund der Nichtlinearität der Ersatzmodellfunktion sowie der Nebenbedingungen wurden im Rahmen dieser Arbeit Optimierungsalgorithmen zur Lösung von Problemen der Form

$$(P) \begin{cases} \min_{x \in \mathbb{R}^n} f(x), \\ \text{sodass} & x_i \in [a_i, b_i] \quad \forall i \in \{1, \dots, n\} \\ \text{und} & h_j(x) \in [c_j, d_j] \quad \forall j \in \{1, \dots, m\} \end{cases}$$

implementiert. Dabei sind die zu minimierende Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ und die Nebenbedingungen $h_j : \mathbb{R}^n \rightarrow \mathbb{R}$ für alle $j \in \{1, \dots, m\}$ nichtlinear. Weiter sind $a_i, c_j \in \mathbb{R} \cup \{-\infty\}$ und $b_i, d_j \in \mathbb{R} \cup \{\infty\}$ für alle $i \in \{1, \dots, n\}$ und $j \in \{1, \dots, m\}$.

In den beiden nachfolgenden Abschnitten werden die zwei verschiedenen, zur Lösung des Problems (P) und somit zur Optimierung der Ersatzmodelle implementierten Algorithmen vorgestellt. Aufgrund der jeweiligen Umformulierung des Problems (P) in Teilprobleme ohne Nebenbedingungen betrachten wir in Abschnitt 4.3 verschiedene Gradienten-basierte Verfahren zur unrestringierten Minimierung. Daraufhin wird die Parallelisierbarkeit der implementierten Algorithmen analysiert und diskutiert.

Alle in diesem Kapitel vorgestellten Verfahren wurden im Rahmen dieser Arbeit implementiert und werden im darauffolgenden Kapitel 5 ausgewertet und verglichen.

4.1 Logarithmische Barriere-Methode

In den 60er Jahren leisteten Fiacco und McCormick [33, 34] wesentliche Beiträge zur Theorie der Barriere- und Penalty-Verfahren für nichtlineare Optimierungsprobleme. Sie führten in diesem Zusammenhang unter anderem den Oberbegriff der Inneren-Punkte-Verfahren ein, zu denen die logarithmische Barriere-Methode gehört (vgl. [35]).

Im Folgenden sei die Problemstellung

$$(PNU) \begin{cases} \min_{x \in D} f(x), \\ \text{sodass} & g_j(x) \leq 0 \quad \forall j \in \{1, \dots, p\} \end{cases}$$

gegeben. Hierbei ist $D \subseteq \mathbb{R}^n$ offen, $f : D \rightarrow \mathbb{R}$ und $g : D \rightarrow \mathbb{R}^p$. Die zulässige Menge sei mit $\mathcal{F} = \{x \in D \mid g_j(x) \leq 0 \quad \forall j \in \{1, \dots, p\}\}$ und die strikt zulässige Menge mit $\mathcal{F}^\circ = \{x \in D \mid g_j(x) < 0 \quad \forall j \in \{1, \dots, p\}\}$ bezeichnet (vgl. [35]). Beachte, dass sich das gegebene zu lösende Problem (P) mit $D = \mathbb{R}^n$, $p = 2n + 2m$ und

$$\begin{aligned} g_i(x) &= a_i - x_i, \\ g_{n+i}(x) &= x_i - b_i, \quad i = 1, \dots, n \\ g_{2n+j}(x) &= c_j - h_j(x), \\ g_{2n+m+j}(x) &= h_j(x) - d_j, \quad j = 1, \dots, m \end{aligned}$$

äquivalent in ein Problem der Form (PNU) umformulieren lässt.

Die Idee der Barriere-Methode besteht in der Betrachtung einer sogenannten logarithmischen Barriere-Funktion für Probleme mit Ungleichungsrestriktionen (vgl. [33, 34]).

Definition 4.1. Sei $r \in \mathbb{R}$, $r > 0$ der Strafparameter, dann ist die logarithmische Barriere-Funktion $\Phi(\cdot, r) : \mathcal{F}^\circ \rightarrow \mathbb{R}$ durch

$$\Phi(x, r) = f(x) - r \sum_{j=1}^p \ln(-g_j(x)) \quad (4.1)$$

definiert.

Zu Beginn des Iterationsverfahrens wählt man nun einen Startparameter r_0 und einen strikt zulässigen Punkt $x^{(0)} \in \mathcal{F}^\circ$. Dieser ist Startpunkt für ein Verfahren zur Berechnung des unrestringierten Minimalpunktes $x(r_0)$ der Funktion $\Phi(\cdot, r_0)$. Die zu diesem Zweck implementierten Verfahren zur unrestringierten nichtlinearen Optimierung werden in Abschnitt 4.3 vorgestellt.

Wegen $-\ln(-g_j(x)) \rightarrow \infty$ für $g_j(x) \nearrow 0$ wird durch den Zusatzterm in (4.1) eine Barriere gegen das Verlassen der zulässigen Menge aufgebaut, sodass wir bei geeigneter Schrittweitenbestimmung $x(r_0) \in \mathcal{F}^\circ$ erhalten. Anschließend wählt man einen Parameter $r_1 < r_0$ und führt die Iteration fort. Wir erhalten das folgende Verfahren (vgl. [35]):

Verfahren 4.1. *Logarithmische Barriere-Methode*

1. Wähle $r_0 = 0.5$, $x^{(0)} \in \mathcal{F}^\circ$ und setze $k = 0$.
2. Berechne einen Minimalpunkt $x(r_k)$ von $\Phi(\cdot, r_k)$ mit einem Verfahren zur unrestringierten Minimierung zum Startpunkt $x^{(0)}$.
Wenn das Abbruchkriterium erfüllt ist, breche die Iteration ab. Die approximierte Lösung ist $x(r_k)$.
3. Setze $r_{k+1} = \frac{1}{2} r_k$.
4. Setze $x^{(0)} = x(r_k)$, $k = k + 1$ und gehe zu Schritt 2.

Für die logarithmische Barriere-Methode wurden zwei verschiedene Abbruchkriterien implementiert. Sei dafür im Folgenden $x(r_k)$ mit $x^{(k)}$ bezeichnet.

Das erste Abbruchkriterium basiert auf der Differenzbetrachtung von aufeinanderfolgenden Funktionswerten und Iterierten (vgl. [35], Bem. 4.2.2). Bei fest vorgegebener Toleranz $\varepsilon > 0$ wird das Verfahren abgebrochen, wenn die Bedingungen

$$\begin{aligned} |f(x^{(k)}) - f(x^{(k-1)})| &\leq \varepsilon, \\ \|x^{(k)} - x^{(k-1)}\| &\leq \sqrt{\varepsilon} \cdot \max \left\{ 1, \|x^{(k)}\| \right\} \quad \text{und} \\ f(x^{(k)}) &\leq f(x^{(k-1)}) \end{aligned}$$

erfüllt sind.

Das zweite Abbruchkriterium beruht hingegen auf einer prozentualen Verbesserung der Funktionswerte und der Nebenbedingungsverletzung. Aus Gründen der Übersichtlichkeit führen wir die prozentualen Abbruchbedingungen an die Nebenbedingungsverletzung schon in diesem Abschnitt ein, obwohl diese aufgrund der strikten Zulässigkeit der Barriere-Methode stets erfüllt sind. Erst im folgenden Abschnitt 4.2, in dem wir dieselben prozentualen Abbruchbedingungen anwenden werden, stellen die Abbruchkriterien an die Neben-

bedingungsverletzung eine tatsächliche Einschränkung dar, da das Augmented-Lagrange-Verfahren eine geringe Unzulässigkeit toleriert.

Sei also $\mathcal{I}_N := \{j \in \{1, \dots, p\} \mid g_j(x) > 0\}$ die Menge der Indizes mit einer Verletzung der Nebenbedingung, so berechnet sich die Nebenbedingungsverletzung $n = n(x)$ durch

$$n(x) = \sum_{j \in \mathcal{I}_N} g_j(x) = \sum_{j=1}^p \max\{0, g_j(x)\}. \quad (4.2)$$

Weiter seien die vier Prozentsätze p_1 bis p_4 fest vorgegeben. Die Barriere-Methode soll abbrechen, sobald eine prozentuale Verbesserung der Startwerte vorliegt, d.h.

$$\begin{aligned} f(x^{(k)}) &\leq p_1 \cdot f(x^{(0)}), \\ n(x^{(k)}) &\leq p_3 \cdot n(x^{(0)}) \end{aligned}$$

und während der letzten Iteration keine signifikante Verbesserung stattgefunden hat, d.h.

$$\begin{aligned} |f(x^{(k)}) - f(x^{(k-1)})| &\leq p_2 \cdot |f(x^{(k-1)})|, \\ |n(x^{(k)}) - n(x^{(k-1)})| &\leq p_4 \cdot n(x^{(k-1)}). \end{aligned}$$

Zudem sollen die Bedingungen $f(x^{(k)}) \leq f(x^{(k-1)})$ und $n(x^{(k)}) \leq n(x^{(k-1)})$ erfüllt sein. Die Art des Abbruchkriteriums wie auch die Werte für Toleranzen respektive Prozentsätze können in einer Eingabedatei festgelegt werden.

4.2 Augmented-Lagrange-Verfahren

Im folgenden Abschnitt wird das Augmented-Lagrange-Verfahren für Ungleichungsrestriktionen vorgestellt. Hierbei handelt es sich um ein Verfahren, welches sequentiell eine um einen quadratischen Strafterm erweiterte Lagrange-Funktion minimiert. Das Verfahren hängt mit dem ursprünglichen quadratischen Penalty-Verfahren zusammen (siehe [36], Abschnitt 17.1), allerdings wird durch die Einführung von approximierten Lagrange-Multiplikatoren die oft auftretende schlechte Konditionierung des quadratischen Penalty-Verfahrens verringert (vgl. [36]).

Sei im Folgenden auch hier ein Problem der Form (PNU) zu lösen. Wir können (PNU) äquivalent zu dem unrestringierten Optimierungsproblem

$$\min_{x \in \mathbb{R}^n} F(x)$$

mit

$$F(x) = \max_{\lambda \geq 0} \left\{ f(x) + \sum_{j=1}^p \lambda_j g_j(x) \right\} = \begin{cases} f(x), & x \in \mathcal{F} \\ \infty, & \text{sonst} \end{cases}$$

umformulieren, wobei die Bedingung $\lambda \geq 0$ komponentenweise zu verstehen ist. Um diese Behauptung zu verifizieren, betrachten wir vorerst den Fall, dass x unzulässig ist. Somit existiert ein $j \in \{1, \dots, p\}$ mit $g_j(x) > 0$. Wählt man nun das zugehörige λ_j beliebig groß und setzt alle anderen $\lambda_i = 0$, so ist $F(x)$ tatsächlich unendlich. Ist x hingegen zulässig, so gilt $g_j(x) \leq 0$ für alle $j \in \{1, \dots, p\}$ und das Maximum wird bei $\lambda = 0$ angenommen. In diesem Fall gilt also $F(x) = f(x)$. Insgesamt folgt somit

$$\min_{x \in \mathbb{R}^n} F(x) = \min_{x \in \mathcal{F}} f(x),$$

also das zu lösende Problem (PNU). Da die Funktion F jedoch Unstetigkeitsstellen beim Übergang zwischen der unzulässigen und der zulässigen Menge aufweist, ist eine direkte Minimierung nicht praktikabel.

Die Lösung dieses Problems besteht in der Approximation der Funktion F durch eine glatte Funktion Φ , welche vom Strafparameter μ_k und vom angenäherten Lagrange-Multiplikator $\lambda^{(k)}$ abhängig ist. Diese Approximation liest sich wie folgt:

$$\Phi(x, \lambda^{(k)}, \mu_k) = \max_{\lambda \geq 0} \left\{ f(x) + \sum_{j=1}^p \lambda_j g_j(x) - \frac{1}{2\mu_k} \sum_{j=1}^p (\lambda_j - \lambda_j^{(k)})^2 \right\}. \quad (4.3)$$

Der hier zusätzlich subtrahierte letzte Term bestraft ein Entfernen des Lagrange-Multiplikators λ von der vorausgegangenen Annäherung $\lambda^{(k)}$. Somit soll sich der neue Maximierer λ nahe $\lambda^{(k)}$ befinden.

Beim Problem (4.3) handelt es sich um eine quadratische Optimierungsaufgabe mit den Nebenbedingungen $\lambda_j \geq 0$, welche explizit gelöst werden kann. Für den Lagrange-Multiplikator erhalten wir somit die Darstellung

$$\lambda_j = \begin{cases} 0, & \text{wenn } \frac{\lambda_j^{(k)}}{\mu_k} + g_j(x) \leq 0 \\ \mu_k g_j(x) + \lambda_j^{(k)} & \text{sonst.} \end{cases} \quad (4.4)$$

Führen wir nun die Indexmenge $\mathcal{I} = \{j \in \{1, \dots, p\} \mid \frac{\lambda_j^{(k)}}{\mu_k} + g_j(x) \leq 0\}$ und deren Komplementmenge \mathcal{I}^C ein, so folgt

$$\Phi(x, \lambda^{(k)}, \mu_k) = f(x) - \sum_{j \in \mathcal{I}} \frac{1}{2\mu_k} \lambda_j^{(k)2} + \sum_{j \in \mathcal{I}^C} \left(\frac{\mu_k}{2} g_j^2(x) + \lambda_j^{(k)} g_j(x) \right).$$

Damit erhalten wir die neue Iterierte $x^{(k+1)}$ als Ergebnis einer unrestringierten Minimierung von $\Phi(x, \lambda^{(k)}, \mu_k)$ bezüglich x . Der aktuelle Lagrange-Multiplikator $\lambda^{(k+1)}$ lässt sich sodann mit Hilfe der Vorschrift (4.4) berechnen. Nutzen wir erneut die Funktion $n = n(x)$ zur Berechnung der Nebenbedingungsverletzung (siehe (4.2)), so erhalten wir den folgenden Algorithmus (vgl. [36], Seite 520 f.):

Verfahren 4.2. *Augmented-Lagrange-Verfahren für (PNU)*

1. Wähle einen Startpunkt $x^{(0)}$ und einen initialen Lagrange-Multiplikator $\lambda^{(0)}$. Setze $k = 0$, $\mu_0 = 10$, $\eta_0 = \mu_0^{-0.1}$ und die Toleranz zur Lösung des unrestringierten Problems $\omega_0 = \mu_0^{-1}$.
2. Berechne einen Minimalpunkt $x^{(k+1)}$ von $\Phi(\cdot, \lambda^{(k)}, \mu_k)$ mit einem Verfahren zur unrestringierten Minimierung zum Startpunkt $x^{(k)}$ mit Abbruchtoleranz ω_k .
3. Falls $n(x^{(k+1)}) \leq \eta_k$, so überprüfe das Abbruchkriterium. Ist dieses erfüllt, so breche die Iteration ab. Die approximierte Lösung ist $x^{(k+1)}$. Ist das Abbruchkriterium nicht erfüllt, so gehe zu Schritt 4.
Falls $n(x^{(k+1)}) > \eta_k$, dann gehe zu Schritt 5.
4. Aktualisiere den Lagrange-Multiplikator nach Vorschrift (4.4). Verkleinere die Toleranzen: $\eta_{k+1} = \frac{\eta_k}{\mu_{k+1}^{0.9}}$, $\omega_{k+1} = \frac{\omega_k}{\mu_{k+1}}$, wobei $\mu_{k+1} = \mu_k$. Setze $k = k + 1$ und gehe zu Schritt 2.

5. Erhöhe den Strafparameter: $\mu_{k+1} = 100\mu_k$. Verkleinere die Toleranzen: $\eta_{k+1} = \mu_{k+1}^{-0.1}$, $\omega_{k+1} = \mu_{k+1}^{-1}$. $\lambda^{(k+1)} = \lambda^{(k)}$. Setze $k = k + 1$ und gehe zu Schritt 2.

Bemerkung 4.1. Sind keine Informationen über den ersten Lagrange-Multiplikator vorhanden, so verwendet man den Nullvektor als initialen Lagrange-Multiplikator $\lambda^{(0)}$.

In Schritt 3 überprüft der Algorithmus, ob sich die Verletzung der Nebenbedingung ausreichend verringert hat. Ist dies der Fall, so ist die Ungleichung

$$n(x^{(k+1)}) \leq \eta_k \quad (4.5)$$

erfüllt. Somit ist eine Vergrößerung des aktuellen Strafparameters nicht notwendig, da schon μ_k eine hinreichend kleine Verletzung der Nebenbedingung realisiert. Der Lagrange-Multiplikator wird nach Vorschrift (4.4) aktualisiert.

Ist im Gegensatz dazu (4.5) nicht erfüllt, so wird der Strafparameter in der darauffolgenden Iteration vergrößert. Das Verfahren legt somit erhöhten Wert auf eine Verkleinerung der Nebenbedingungsverletzung in der nächsten Iteration. Es wird auf eine Aktualisierung des Lagrange-Multiplikators verzichtet; das Hauptaugenmerk liegt in der Verbesserung der Zulässigkeit (vgl. [36]).

Auch für das Augmented-Lagrange-Verfahren wurden zwei verschiedene Abbruchkriterien implementiert. Das erste Abbruchkriterium fordert eine hinreichend kleine Verletzung der Nebenbedingung sowie eine hinreichend kleine Norm des Gradienten der zu minimierenden Funktion Φ (vgl. [36]). Hierbei wird die zweite Bedingung (4.6) aufgrund der Verwendung von Gradienten-basierten Verfahren zur unrestringierten Minimierung von Φ gefordert. Sei also die Toleranz $\varepsilon > 0$ fest vorgegeben, so bricht das Augmented-Lagrange-Verfahren ab, sobald die Ungleichungen

$$\begin{aligned} n(x^{(k+1)}) &\leq \sqrt{\varepsilon} \quad \text{und} \\ \|\nabla\Phi(x^{(k+1)}, \lambda^{(k)}, \mu_k)\| &\leq \varepsilon \end{aligned} \quad (4.6)$$

erfüllt sind. Das zweite Abbruchkriterium beruht wiederum auf einer prozentualen Verbesserung der Funktionswerte und der Nebenbedingungsverletzung und ist analog zu dem der logarithmischen Barriere-Methode.

4.3 Verfahren zur unrestringierten Optimierung

Im Verlauf der vorherigen Abschnitte haben wir gesehen, dass beide Verfahren, die logarithmische Barriere-Methode wie auch das Augmented-Lagrange-Verfahren, während des Optimierungsprozesses ein Verfahren zur unrestringierten Minimierung einer reellwertigen, differenzierbaren Funktion $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}$ aufrufen. Im Rahmen dieser Arbeit wurden vier verschiedene Gradienten-basierte Verfahren implementiert, die im Folgenden kurz vorgestellt werden sollen. Unterschiede und Vorteile der jeweiligen Verfahren werden untersucht und herausgestellt.

4.3.1 Gradientenverfahren

Beim Gradientenverfahren oder auch Verfahren des steilsten Abstiegs handelt es sich um ein Minimierungsverfahren, welches als Abstiegs- bzw. Suchrichtung $d^{(k)}$ den Antigradienten $-\nabla\Phi(x^{(k)})$ verwendet (vgl. [35]).

Verfahren 4.3. *Gradientenverfahren*

1. Wähle einen Startpunkt $x^{(0)}$ und setze $k = 0$.
2. Ist das Abbruchkriterium erfüllt, so stoppe das Verfahren. Die approximierte Lösung ist $x^{(k)}$.
3. Berechne eine effiziente Schrittweite σ_k [¶] und die Abstiegsrichtung $d^{(k)} = -\nabla\Phi(x^{(k)})$, dann setze $x^{(k+1)} = x^{(k)} + \sigma_k d^{(k)}$.
4. Setze $k = k + 1$ und gehe zu Schritt 2.

Es wurden zwei Abbruchkriterien implementiert. Das erste Abbruchkriterium basiert auf der Grundidee, dass der Gradient der zu minimierenden Funktion im Optimum Null ist. Somit bricht das Gradientenverfahren ab, sobald bei fest vorgegebener Toleranz $\omega > 0$ die Ungleichung

$$\|\nabla\Phi(x^{(k)})\| \leq \omega \cdot \max\{1, \|x^{(k)}\|\}$$

erfüllt ist (vgl. [35], Seite 69). Zusätzlich wurde noch ein zweites Abbruchkriterium basierend auf dem Abstand zweier Iterierter $x^{(k)}$ und $x^{(k+1)}$ implementiert. Das Verfahren soll abbrechen, wenn für festes $\omega_x > 0$ die Bedingung

$$\max_{i=1,\dots,n} \frac{|x_i^{(k+1)} - x_i^{(k)}|}{\max\{1, x_i^{(k+1)}\}} \leq \omega_x \quad (4.7)$$

erfüllt ist.

Konvergenz des Gradientenverfahrens

Sei im Folgenden die Niveaumenge $\mathcal{N}_0 = \{x \in \mathbb{R}^n \mid \Phi(x) \leq \Phi(x^{(0)})\}$ nichtleer. Weiter sei Φ zweimal stetig differenzierbar auf einer offenen, konvexen Obermenge von \mathcal{N}_0 und die Hessematrix von Φ gleichmäßig positiv definit auf dieser Obermenge^{||}. Unter den eben genannten Voraussetzungen lässt sich bei Verwendung von sogenannten effizienten Schrittweiten (vgl. Abschnitt 4.3.5) die lineare Konvergenz des Verfahrens gegen den eindeutig bestimmten, globalen Minimalpunkt \tilde{x} von Φ nachweisen (vgl. [35], Satz 4.2.42).

Da wir bei Optimierung der Ersatzmodelle die eben genannten Voraussetzungen nicht garantieren können, ist dieses Konvergenzresultat nur von theoretischem Interesse. Es zeigt jedoch, dass das Gradientenverfahren im Vergleich zu den im Folgenden betrachteten Verfahren theoretisch das schlechteste Konvergenzverhalten aufweist. Begründen lässt sich dieses Verhalten aufgrund der bei Verwendung effizienter Schrittweiten „annähernd orthogonalen“ Suchrichtungen $d^{(k)}$ des Gradientenverfahrens. Das daraus resultierende Zick-Zack-Verhalten nahe des Minimums hat eine meist langsame Konvergenz des Verfahrens zur Folge (vgl. [35], S. 99).

[¶]siehe Abschnitt 4.3.5 und Definition 4.2 der effizienten Schrittweite, Seite 45

^{||}Die Hessematrix Φ'' ist gleichmäßig positiv definit auf $\Omega \supset \mathcal{N}_0$, wenn eine Konstante $C > 0$ existiert, sodass für alle $x \in \Omega$ und alle $d \in \mathbb{R}^n$ die Ungleichung $d^T \Phi''(x) d \geq C \|d\|^2$ gilt.

4.3.2 Fletcher-Reeves-Verfahren

Das Fletcher-Reeves-Verfahren stellt eine Erweiterung des Verfahrens der konjugierten Gradienten für quadratische Zielfunktionen (kurz: CG-Verfahren, siehe [35], S. 75 ff.) dar. Es wurde für allgemeine nichtlineare Probleme entwickelt und erstmals von Fletcher und Reeves [37] untersucht (vgl. [35]).

In der ersten Iteration startet das Fletcher-Reeves-Verfahren wie das Gradientenverfahren in Richtung des steilsten Abstiegs. In den darauffolgenden Iterationen wird jedoch ein Anteil der Suchrichtung der vorherigen Iteration zum Antigradienten hinzuaddiert (vgl. [38]). Wir erhalten das folgende Verfahren:

Verfahren 4.4. Fletcher-Reeves-Verfahren

1. Wähle einen Startpunkt $x^{(0)}$ und setze $k = 0$, $d^{(0)} = -\nabla\Phi(x^{(0)})$.
2. Ist das Abbruchkriterium erfüllt, so stoppe das Verfahren. Die approximierte Lösung ist $x^{(k)}$.
3. Berechne eine effiziente Schrittweite σ_k und setze $x^{(k+1)} = x^{(k)} + \sigma_k d^{(k)}$.
4. Berechne den Parameter $\beta_k = \frac{\|\nabla\Phi(x^{(k+1)})\|^2}{\|\nabla\Phi(x^{(k)})\|^2}$ und die konjugierte Abstiegsrichtung $d^{(k+1)} = -\nabla\Phi(x^{(k+1)}) + \beta_k d^{(k)}$.
5. Setze $k = k + 1$ und gehe zu Schritt 2.

Für das Fletcher-Reeves-Verfahren wurden ebenso zwei unterschiedliche Abbruchkriterien implementiert. Diese sind analog zu denen des Gradientenverfahrens (siehe Abschnitt 4.3.1).

Bemerkung 4.2. In der Praxis erzielt das Verfahren oft nach einer gewissen Anzahl an Iterationen keinen signifikanten Fortschritt mehr. Es hat sich gezeigt, dass in diesem Fall ein sogenannter Restart, also ein Neustart des Verfahrens am aktuellen Iterationspunkt, hilfreich sein kann. Üblicherweise führt man diesen nach je n Iterationen durch (vgl. [39]).

Konvergenz des Fletcher-Reeves-Verfahrens

Unter denselben Voraussetzungen an die Niveaumenge, die Differenzierbarkeit und die zweite Ableitung, wie wir sie zuvor bei der Konvergenzbetrachtung des Gradientenverfahrens gefordert haben, lässt sich bei Verwendung von exakten Schrittweiten eine Konvergenz der $(x^{(k)})_k$ gegen den eindeutig bestimmten, globalen Minimalpunkt \tilde{x} zeigen (vgl. [35]). Im Fall der Restart-Variante kann zudem eine sogenannte n -Schritt-quadratische Konvergenz des Verfahrens nachgewiesen werden, d.h. es gilt

$$\|x^{(n(k+1))} - \tilde{x}\| \leq C \|x^{(nk)} - \tilde{x}\|^2$$

mit einer von k unabhängigen Konstante $C > 0$ (siehe [40]). Der Fehler quadriert sich somit nach n Iterationen.

4.3.3 Polak-Ribière-Verfahren

Das Polak-Ribière-Verfahren ist ebenfalls eine Erweiterung des CG-Verfahrens für allgemeine nichtlineare Zielfunktionen. Der Unterschied zum Fletcher-Reeves-Verfahren 4.4 liegt in der Wahl der β_k in Schritt 4. Polak und Ribière schlagen in [41]

$$\beta_k = \frac{\|\nabla\Phi(x^{(k+1)})\|^2 - \nabla\Phi(x^{(k+1)})^T \nabla\Phi(x^{(k)})}{\|\nabla\Phi(x^{(k)})\|^2}$$

vor und zeigen im Anschluss die Konvergenz des Verfahrens.

Auch bei diesem Verfahren kann eine Ineffektivität der Suchrichtungen nach n Iterationen auftreten, allerdings zeigt das Polak-Ribière-Verfahren für Iterationen über n hinaus im Allgemeinen ein besseres Verhalten als das Fletcher-Reeves-Verfahren. Zudem erhält man grundsätzlich ein besseres numerisches Verhalten mit β_k nach Polak und Ribière (vgl. [38]).

4.3.4 Quasi-Newton-Verfahren

Algorithmen wie das Quasi-Newton-Verfahren, welche eine der Newton-Richtung (vgl. [36]) $d^{(k)} = -\Phi''(x^{(k)})^{-1} \nabla\Phi(x^{(k)})$ ähnliche Suchrichtung verwenden, neigen zu einem besseren Konvergenzverhalten und einer geringeren Iterationsanzahl als die zuvor besprochenen Gradienten-basierten Verfahren. Aus diesem Grund wurde das Quasi-Newton-Verfahren ebenfalls im Rahmen dieser Arbeit implementiert und ausgewertet und soll im Folgenden näher betrachtet werden.

Beim Quasi-Newton-Verfahren wird eine Suchrichtung der Form

$$x - x^{(k)} = d^{(k)} = -A^{-1} \nabla\Phi(x^{(k)}) \quad (4.8)$$

verwendet. Ist die Matrix A positiv definit, so handelt es sich bei $d^{(k)}$ um eine Abstiegsrichtung (vgl. [35]). Die Idee des Quasi-Newton-Verfahrens besteht nun in der Approximation der inversen Hessematrix $(\Phi''(x^{(k)}))^{-1}$ durch eine einfach zu berechnende Matrix $B^{(k)} = A^{-1}$. Insbesondere soll der Übergang von $B^{(k)}$ zu $B^{(k+1)}$, das sogenannte Update, ohne großen Aufwand möglich sein (vgl. [35, 42]).

Bemerkung 4.3. Die Verwendung einer approximierten Hessematrix hat einen entscheidenden Vorteil gegenüber dem nur lokal konvergenten Newton-Verfahren. Befinden wir uns weit vom Minimum entfernt, so ist die positive Definitheit der Hessematrix nicht mehr sichergestellt und das Newton-Verfahren entfernt sich womöglich vom Minimum. Das Quasi-Newton-Verfahren startet jedoch mit einer positiv definiten, symmetrischen Matrix $A^{(0)}$, meist der Einheitsmatrix I , und behält diese Eigenschaften aufgrund der Art des Updates für alle folgenden Matrizen $A^{(k)}$ bei. Befinden wir uns also weit vom Minimum entfernt, so ist sichergestellt, dass $d^{(k)}$ eine Abstiegsrichtung ist. Nähern wir uns dem Minimum, so approximieren wir die tatsächliche Hesse-Matrix und das Quasi-Newton-Verfahren zeigt ein gutes Konvergenzverhalten (vgl. [42]).

Im Folgenden skizzieren wir kurz die Herleitung der bekanntesten Update-Formel für die approximierte Inverse $B^{(k+1)}$ der Hessematrix nach Broyden, Fletcher, Goldfarb und Shanno (vgl. [36]). Eine exakte Herleitung dieser sogenannten BFGS-Update-Formel findet sich beispielsweise in Polak [43].

Mit Hilfe von Gleichung (4.8) für k und $k+1$ sowie anschließender Subtraktion der beiden Gleichungen folgt

$$x^{(k+1)} - x^{(k)} = A^{-1} \left(\nabla \Phi(x^{(k+1)}) - \nabla \Phi(x^{(k)}) \right). \quad (4.9)$$

Folglich verlangen wir von der neuen Iterierten $B^{(k+1)}$, dass diese die Gleichung (4.9) erfüllt, d.h.

$$x^{(k+1)} - x^{(k)} = B^{(k+1)} \left(\nabla \Phi(x^{(k+1)}) - \nabla \Phi(x^{(k)}) \right).$$

Des Weiteren soll $B^{(k+1)}$ als Annäherung der Hessematrix symmetrisch und positiv definit sein und sich in gewisser Weise nahe der vorherigen Matrix $B^{(k)}$ befinden, d.h. $B^{(k+1)}$ ist die Lösung von

$$(*) \left\{ \begin{array}{l} \min_{B \in \mathbb{R}^{n \times n}} \|B - B^{(k)}\|, \\ \text{sodass } B = B^T, \\ x^{(k+1)} - x^{(k)} = B \left(\nabla \Phi(x^{(k+1)}) - \nabla \Phi(x^{(k)}) \right) \text{ und} \\ x^T B x > 0 \quad \forall x \in \mathbb{R}^n. \end{array} \right.$$

Wählen wir für die Matrix-Norm in $(*)$ die gewichtete Frobenius-Norm (vgl. [36], S. 139), so erhalten wir durch Lösung von $(*)$ die BFGS-Update-Formel für $B^{(k+1)}$ (vgl. [42])

$$\begin{aligned} B^{(k+1)} = & B^{(k)} + \frac{(x^{(k+1)} - x^{(k)}) \otimes (x^{(k+1)} - x^{(k)})}{(x^{(k+1)} - x^{(k)}) \cdot (\nabla \Phi(x^{(k+1)}) - \nabla \Phi(x^{(k)}))} \\ & - \frac{[B^{(k)} (\nabla \Phi(x^{(k+1)}) - \nabla \Phi(x^{(k)}))] \otimes [B^{(k)} (\nabla \Phi(x^{(k+1)}) - \nabla \Phi(x^{(k)}))]}{(\nabla \Phi(x^{(k+1)}) - \nabla \Phi(x^{(k)})) \cdot B^{(k)} (\nabla \Phi(x^{(k+1)}) - \nabla \Phi(x^{(k)}))} \\ & + [(\nabla \Phi(x^{(k+1)}) - \nabla \Phi(x^{(k)})) \cdot B^{(k)} (\nabla \Phi(x^{(k+1)}) - \nabla \Phi(x^{(k)}))] u \otimes u, \end{aligned} \quad (4.10)$$

mit u definiert als der Vektor

$$\begin{aligned} u = & \frac{(x^{(k+1)} - x^{(k)})}{(x^{(k+1)} - x^{(k)}) \cdot (\nabla \Phi(x^{(k+1)}) - \nabla \Phi(x^{(k)}))} \\ & - \frac{B^{(k)} \cdot (\nabla \Phi(x^{(k+1)}) - \nabla \Phi(x^{(k)}))}{(\nabla \Phi(x^{(k+1)}) - \nabla \Phi(x^{(k)})) \cdot B^{(k)} (\nabla \Phi(x^{(k+1)}) - \nabla \Phi(x^{(k)}))}. \end{aligned}$$

Hierbei bezeichnet \otimes das dyadische Produkt zweier Vektoren. Somit erhalten wir das folgende Verfahren.

Verfahren 4.5. *Quasi-Newton-Verfahren mit BFGS-Update*

1. Wähle einen Startpunkt $x^{(0)}$ und setze $k = 0$, $B^{(0)} = I$.
2. Ist das Abbruchkriterium erfüllt, so stoppe das Verfahren. Die approximierte Lösung ist $x^{(k)}$.
3. Berechne für $k > 0$ die Matrix $B^{(k)}$ nach (4.10), $d^{(k)} = -B^{(k)} \nabla \Phi(x^{(k)})$.

4. Berechne eine effiziente Schrittweite σ_k und setze $x^{(k+1)} = x^{(k)} + \sigma_k d^{(k)}$.
5. Setze $k = k + 1$ und gehe zu Schritt 2.

Bemerkung 4.4. Ein Nachteil des Quasi-Newton-Verfahrens besteht in der Notwendigkeit der Speicherung der in der Regel vollbesetzten Matrizen $B^{(k)}$ (vgl. [35]). Im Gegensatz zu den drei vorherigen Gradienten-basierten Verfahren, welche einen Speicherplatz der Größenordnung n benötigen, benötigt dieses Verfahren einen der Größenordnung n^2 . Für kleine n ist dieser Nachteil jedoch vernachlässigbar.

Ebenso wie für die vorherigen Verfahren wurden für das Quasi-Newton-Verfahren zwei verschiedene Abbruchkriterien implementiert. Die Bedingung (4.7) an die Differenz zweier aufeinanderfolgender Iterierter wurde vom Gradientenverfahren übernommen. Einzig die Bedingung an den Gradienten wurde variiert. Das Quasi-Newton-Verfahren soll abbrechen, sobald bei fest vorgegebener Toleranz $\omega > 0$ die Bedingung

$$\max_{i=1,\dots,n} \frac{|\nabla \Phi(x^{(k+1)})_i| \cdot \max\{|x_i^{(k+1)}|, 1\}}{\max\{1, \Phi(x^{(k)})\}} \leq \omega$$

erfüllt ist (siehe [42]).

Konvergenz des Quasi-Newton-Verfahrens mit BFGS-Update

Nocedal und Wright zeigen in [36] die globale lineare Konvergenz des Quasi-Newton-Verfahrens mit BFGS-Update unter ähnlichen Voraussetzungen wie wir sie zuvor für die Konvergenz des Gradientenverfahrens benötigt haben. Hier wird die zweimal stetige Differenzierbarkeit von Φ sowie die Konvexität der Niveaumenge \mathcal{N}_0 gefordert. Zusätzlich gelte mit den Konstanten $m > 0$ und $M > 0$ die Bedingung

$$m\|z\|^2 \leq z^T \Phi''(x)z \leq M\|z\|^2$$

für alle $z \in \mathbb{R}^n$ und $x \in \mathcal{N}_0$. Dies sichert unter anderem die gleichmäßige positive Definitheit der Hessematrix von Φ . Bei Verwendung effizienter Schrittweiten lässt sich unter diesen Voraussetzungen die lineare Konvergenz des Verfahrens gegen den eindeutig bestimmten Minimalpunkt \tilde{x} von Φ nachweisen (siehe [36], S. 172 ff.). Man kann sogar zeigen, dass die Folge $(\|x^{(k)} - \tilde{x}\|)_k$ so schnell gegen Null konvergiert, dass

$$\sum_{k=1}^{\infty} \|x^{(k)} - \tilde{x}\| < \infty.$$

Dies nutzen Nocedal und Wright, um unter der zusätzlichen Voraussetzung der Lipschitz-Stetigkeit der Hessematrix bei \tilde{x}^{**} die superlineare Konvergenz des BFGS-Verfahrens zu zeigen (siehe [36], Theorem 6.6). Hierbei handelt es sich um die Charakterisierung der superlinearen Konvergenz nach Dennis und Moré, d.h.

$$\lim_{k \rightarrow \infty} \frac{\|((B^{(k)})^{-1} - \Phi''(\tilde{x}))(x^{(k+1)} - x^{(k)})\|}{\|x^{(k+1)} - x^{(k)}\|} = 0.$$

**Die Hessematrix Φ'' ist Lipschitz-stetig bei \tilde{x} , wenn eine Konstante $L > 0$ existiert, sodass $\|\Phi''(x) - \Phi''(\tilde{x})\| \leq L\|x - \tilde{x}\|$ für alle x in einer Umgebung von \tilde{x} gilt.

4.3.5 Schrittweitenbestimmung mit dem Armijo-Verfahren

Im Folgenden sei die Niveaumenge \mathcal{N}_0 kompakt und Φ auf einer konvexen, offenen Obermenge von \mathcal{N}_0 stetig differenzierbar. Weiter sei die aktuelle Iterierte mit $x^{(k)}$ bezeichnet und $d^{(k)}$ eine Abstiegsrichtung. Wie Beispiel 4.2.18 in [35] zeigt, ist es nicht ausreichend, eine hinreichend kleine Schrittweite σ_k mit $\Phi(x^{(k)} + \sigma_k d^{(k)}) < \Phi(x^{(k)})$ zu wählen, um Konvergenz zu erhalten. Es kann der Fall auftreten, dass die Folge der $(x^{(k)})_k$ nicht gegen den Minimalpunkt \tilde{x} konvergiert. Aus diesem Grund ist es notwendig, eine sogenannte effiziente Schrittweitenstrategie zu wählen (vgl. [35], S. 79 f.).

Definition 4.2. Eine Schrittweite σ_k erfüllt das Prinzip des hinreichenden Abstiegs, falls es von $x^{(k)} \in \mathcal{N}_0$ und $d^{(k)}$ unabhängige positive Konstanten c_1 und $c_2 \in \mathbb{R}$ gibt so, dass

$$\Phi(x^{(k)} + \sigma_k d^{(k)}) \leq \Phi(x^{(k)}) + c_1 \sigma_k \nabla \Phi(x^{(k)})^T d^{(k)} \quad (4.11)$$

$$\text{und } \sigma_k \geq -c_2 \frac{\nabla \Phi(x^{(k)})^T d^{(k)}}{\|d^{(k)}\|^2}. \quad (4.12)$$

Eine Schrittweite σ_k heißt effizient, wenn es eine von $x^{(k)} \in \mathcal{N}_0$ und $d^{(k)}$ unabhängige positive Konstante $c \in \mathbb{R}$ gibt so, dass

$$\Phi(x^{(k)} + \sigma_k d^{(k)}) \leq \Phi(x^{(k)}) - c \left(\frac{\nabla \Phi(x^{(k)})^T d^{(k)}}{\|d^{(k)}\|} \right)^2.$$

Bemerkung 4.5. Eine Schrittweitenstrategie, die das Prinzip des hinreichenden Abstiegs erfüllt, ist auch effizient. Verwendet man eine effiziente Schrittweite, so kann man

$$\frac{\nabla \Phi(x^{(k)})^T d^{(k)}}{\|d^{(k)}\|} \rightarrow 0$$

für $k \rightarrow \infty$ zeigen. Fordert man zusätzlich gradientenbezogene Suchrichtungen^{††} $d^{(k)}$, so kann man $\nabla \Phi(x^{(k)}) \rightarrow 0$ für $k \rightarrow \infty$ nachweisen (siehe [35], Satz 4.2.25). Die in den Verfahren 4.3 bis 4.5 gewählten Suchrichtungen sind gradientenbezogen (vgl. [35], S. 84 f.).

Das im Rahmen dieser Arbeit implementierte Armijo-Verfahren zur numerischen Schrittweitenberechnung basiert auf den Arbeiten von Goldstein [44] und Armijo [45], weshalb es auch oft Goldstein-Armijo-Verfahren genannt wird (vgl. [35]). Die Schrittweite wird im Verlauf des Verfahrens sukzessive verkleinert, bis die Abstiegsbedingung (4.11) für $c_1 = \delta$ mit $\delta \in]0, 1[$ erfüllt ist. Damit die Schrittweiten nicht zu klein werden, fordert man zusätzlich (4.12) mit einer von der Abstiegsrichtung unabhängigen Konstante $c_2 > 0$. Im Folgenden bezeichnen wir die aktuellen Iterierten $x^{(k)}$ und $d^{(k)}$ der Einfachheit halber mit x respektive d . Im Armijo-Verfahren benötigen wir des Weiteren die von x und d unabhängigen Konstanten $\gamma > 0$ und $0 < \beta_1 \leq \beta_2 < 1$ (vgl. [35]).

^{††}Die Folge der Abstiegsrichtungen $(d^{(k)})_k$ heißt gradientenbezogen, falls eine von k unabhängige Konstante $c_3 > 0$ existiert, sodass $-\nabla \Phi(x^{(k)})^T d^{(k)} \geq c_3 \|\nabla \Phi(x^{(k)})\| \|d^{(k)}\| \forall k \geq 0$ gilt (vgl. [35]).

Verfahren 4.6. *Armijo-Verfahren*

1. Wähle eine Startschrittweite σ_0 für die (4.12) mit $c_2 = \gamma$ gilt und setze $j = 0$.
2. Ist (4.11) mit $c_1 = \delta$ für σ_j erfüllt, so stoppe das Verfahren; σ_j ist eine effiziente Schrittweite.
3. Wähle $\sigma_{j+1} \in [\beta_1 \sigma_j, \beta_2 \sigma_j]$.
4. Setze $j = j + 1$ und gehe zu Schritt 2.

Das Verfahren gibt keine konkreten Werte für die verwendeten Konstanten an, allerdings haben sich die folgenden Werte bewährt (siehe [35], S. 92 ff.):

$$\begin{aligned}\delta &= 0.01, \\ \gamma &= 10^{-4} \quad \text{und} \\ \beta_1 &= \beta_2 = \beta = \frac{1}{2}.\end{aligned}$$

Ferner sollte eine sinnvolle Startschrittweite gewählt werden. Hier besteht die Grundidee darin, die Startschrittweite σ_0 als Approximation der optimalen Schrittweite zu wählen. Letztendlich erhält man nach weiterführenden Überlegungen die folgende Formel zur Berechnung der Startschrittweite $\tilde{\sigma}_0$ (vgl. [35]):

$$\tilde{\sigma}_0 = \min \left\{ -\tilde{\gamma} \frac{\nabla \Phi(x)^T d}{\|d\|^2}, \max \left\{ -\gamma \frac{\nabla \Phi(x)^T d}{\|d\|^2}, \sigma_0 \right\} \right\} \quad (4.13)$$

mit

$$\sigma_0 = -\frac{\nabla \Phi(x)^T d}{2(\Phi(x+d) - \Phi(x) - \nabla \Phi(x)^T d)}.$$

Hierbei hat sich $\tilde{\gamma} = 10^4$ praktisch bewährt (vgl. [35]).

Während der im Rahmen der Arbeit durchgeführten Tests der einzelnen Verfahren hat sich herausgestellt, dass die Konvergenzgeschwindigkeit unter anderem stark von der Startschrittweite abhängig ist. Das Armijo-Schrittweitenverfahren kann diese nur verkleinern (vgl. Verfahren 4.6, Schritt 3). Ist die Startschrittweite nun über viele Iterationen aufgrund der nur heuristischen Formel (4.13) zu klein gewählt, so kann sich die benötigte Anzahl an Iterationen und somit auch die Konvergenzgeschwindigkeit unverhältnismäßig stark verschlechtern (siehe auch Kapitel 5). In der Implementation des Verfahrens gibt es daher die Möglichkeit eine untere Grenze für die Startschrittweite (in der Eingabedatei mit SIGMA_SHORT bezeichnet, vgl. Anhang B.1, S. 100) anzugeben. Unterschreitet $\tilde{\sigma}_0$ diese Grenze, so wird die Startschrittweite automatisch auf den Wert 1.5 gesetzt. Während der Tests hat sich in vielen Fällen SIGMA_SHORT = 0.001 bewährt, allerdings ist eine Anpassung dieser Grenze auch bei laufender Optimierung über die Eingabedatei durch den Benutzer möglich.

4.4 Parallelisierungsaspekte

Die Parallelisierung eines Algorithmus ist grundsätzlich immer dann sinnvoll, wenn viele gleichgeartete unabhängige Berechnungen durchgeführt werden und diese Berechnungen relativ viel Zeit in Anspruch nehmen.

Im Fall unserer Gradienten-basierten Optimierungsalgorithmen bietet die Schrittweitensteuerung das beste Potential zur Parallelisierung. Warum dies sinnvoll ist und wie diese Parallelisierung praktisch vorgenommen werden kann, soll im nächsten Unterabschnitt analysiert werden. Um einen Eindruck davon zu gewinnen, inwieweit Parallelisierung bei der Verdichter-Optimierung im DLR eine Rolle spielt, wollen wir im darauffolgenden Abschnitt kurz auf die schon implementierten Parallelisierungen in AutoOpti und TRACE eingehen.

4.4.1 Parallelisierbarkeit der Schrittweitensteuerung

Betrachtet man die implementierten Optimierungsalgorithmen unabhängig davon, ob es sich dabei um das Augmented-Lagrange-Verfahren, die Barriere-Methode oder eines der Verfahren zur unrestringierten Optimierung handelt, so stellt man schnell fest, dass diese größtenteils auf Operationen beruhen, deren Ausführung sequentiell erfolgen muss. Die einzig sinnvolle Möglichkeit, eine Parallelisierung einzubauen, bietet die Schrittweitenberechnung durch das Armijo-Verfahren (vgl. Abschnitt 4.3.5). Genauer gesagt handelt es sich um eine Parallelisierung der hier durchgeführten *Line Search*, also bei gegebenem $x^{(k)}$ und zugehöriger Suchrichtung $d^{(k)}$ das Suchen einer effizienten Schrittweite durch sukzessives Verkleinern von σ entlang der Linie $\{x^{(k)} + \sigma d^{(k)} \mid \sigma \in]0, \tilde{\sigma}_0]\}$.

Bei erneuter Betrachtung des Armijo-Verfahrens 4.6 wird ersichtlich, dass in Schritt 2 aufgrund der Überprüfung der Ungleichung (4.11) in jeder Armijo-Iteration der Term

$$\Phi(x^{(k)} + \sigma_j d^{(k)}) \quad (4.14)$$

für die zu überprüfende Schrittweite σ_j ausgewertet werden muss. Ist Ungleichung (4.11) nicht erfüllt, so wird σ_j entsprechend der Vorschrift verkleinert und die Gültigkeit der Ungleichung wird erneut überprüft.

Im Allgemeinen sollte die Anzahl der Iterationen des Armijo-Verfahrens eine geringe Größenordnung aufweisen. Bei einer schlechten Wahl der Startschrittweite kann jedoch der Fall auftreten, dass der Term (4.14) bis zu dreißigmal, wenn nicht noch öfter, ausgewertet werden muss. Aufgrund der fixen Vorschrift zur Verkleinerung der Schrittweite - in unserer Implementierung wird diese in jeder Iteration halbiert - ist es möglich, eine feste Anzahl von Auswertungen des Terms (4.14), seien es zehn oder auch dreißig, parallel durchzuführen. Sei $q \in \mathbb{N}$ diese feste Anzahl an Auswertungen, so berechnen wir den Term (4.14) für alle

$$\sigma_j = \left(\frac{1}{2}\right)^q \tilde{\sigma}_0$$

mit $j = 0, \dots, q - 1$ parallel. Im Anschluss können wir sogleich überprüfen, welche der σ_j die Ungleichung (4.11) erfüllen. Trifft dies auf mehrere σ_j zu, so wählen wir die Schrittweite mit dem kleinsten Funktionswert $\Phi(x^{(k)} + \sigma_j d^{(k)})$. Erfüllt keines der σ_j die Ungleichung (4.11), so führen wir den Prozess weiter fort.

Eine Möglichkeit, die oben besprochene Parallelisierung in der Praxis zu implementieren, bietet das Message Passing Interface (vgl. Anhang A.2, Seite 97).

Bemerkung 4.6. Im Fall einer Optimierung der Ersatzmodelle, wie sie im Rahmen dieser Arbeit durchgeführt wird, mag der Zeitgewinn den Aufwand der Parallelisierung nicht aufwiegen. Plant man allerdings, wie in Abschnitt 2.2.1 auf Seite 8 angedeutet, eine Formoptimierung der richtigen Modelle durchzuführen (vgl. Herangehensweise Abschnitt 1.1),

so zieht eine Funktionsauswertung von Φ eine Reihe von Prozessen nach sich, unter anderem die äußerst kostspieligen CFD- und CSM-Simulationen (siehe Abschnitt 2.2.2). Um in diesem Fall die sowieso schon extrem langen Rechenzeiten zu verringern, ist eine Parallelisierung der Line Search unerlässlich.

4.4.2 Parallelisierung in AutoOpti und TRACE

Wie bereits in Abschnitt 2.2 angesprochen, handelt es sich bei AutoOpti um ein hochparalleles Programm zur Optimierung. Die Parallelisierung wurde im Hauptoptimierungsprozess (siehe Abb. 2.3, rechte Seite) wie auch im Ersatzmodell-Prozess (linke Seite der Abb. 2.3) vorgenommen.

Im Hauptprozess dient die Parallelisierung der nebenläufigen Fitness-Berechnung von Individuen respektive Verdichterkonfigurationen, die zuvor im genetischen Algorithmus oder mit Hilfe der Ersatzmodelle generiert wurden. Diese parallele Berechnung wird in den sogenannten Opti-Slave-Prozessen (vgl. Abb. 2.3) vorgenommen. Die Fitness-Berechnung ist aufgrund des Durchlaufs der gesamten Design-Prozesskette ein sehr zeitaufwendiger Prozess, sodass eine Parallelisierung hier unverzichtbar ist.

Auch im Bereich der Ersatzmodelle wurde eine Parallelisierung implementiert. Bevor die Optimierung auf den Ersatzmodellen starten kann, müssen diese trainiert, also generiert werden (vgl. Abschnitt 3.2.1, Seite 22). In AutoOpti wird eine feste Anzahl N Ersatzmodelle parallel trainiert. Dies geschieht mit Hilfe der Meta-Slave-Prozesse (vgl. Abb. 2.3). Im Anschluss wird aus den N Ersatzmodellen ein sogenanntes durchschnittliches Ersatzmodell (engl. *average metamodel*) generiert. Dieses Ersatzmodell, welches wir auch im Rahmen der Arbeit betrachten, wird sodann optimiert (siehe [1]).

AutoOpti erlangt ein weiteres Level der Parallelisierung aufgrund des hochparallelen Strömungslösers TRACE, welcher im Laufe der Design-Prozesskette zur CFD-Simulation genutzt wird (vgl. Abschnitt 2.2.2, Seite 10). TRACE selbst ist wiederum auf mehreren Ebenen parallelisiert. Auf der höchsten Ebene erfolgt die Parallelisierung mittels MPI. Ein Level niedriger ist auf Ebene der Mehrkernprozessoren eine OpenMP-Parallelisierung implementiert. Insgesamt wird also eine hybride OpenMP-MPI-Parallelisierungsstrategie verwendet (vgl. [46]). Zur Erläuterung der in diesem Abschnitt verwendeten Begriffe sei auf Anhang A.2 verwiesen.

In TRACE wird das Rechengebiet - in unserem Fall alle durchströmten Bereiche des Verdichters - in unterschiedlich große Blöcke unterteilt (vgl. Abb. 2.4, S. 10). Die Form und Größe dieser Blöcke orientiert sich an den Designelementen. Beispielsweise muss die Grenzschicht^{††} aufgrund der dort vermehrt auftretenden turbulenten Vorgänge genauer aufgelöst beziehungsweise feiner diskretisiert werden als weiter außen liegende Bereiche.

Die einzelnen Blöcke können nun je nach Aufgabenstellung sowohl strukturiert als auch unstrukturiert diskretisiert werden. TRACE bietet hier die Möglichkeit auf beiden Gitterformen zu rechnen. Auf dem so erzeugten Gitter kann nun die RANS-Gleichung (2.1) in diskretisierter Form mit Hilfe von MPI parallel gelöst werden. Die erforderliche Kommunikation aufgrund der sich überlappenden Ränder der Blöcke erfolgt mittels MPI über Nachrichten.

^{††}Bereich des strömenden Gases, welcher sich in Wandnähe, also in unserem Fall nahe des Verdichterprofils, befindet; für eine genaue Definition sei auf [5] verwiesen.

Für die Parallelisierung ist die stark differierende Größe der Blöcke problematisch (siehe bspw. Abb. 2.4). Die Lastverteilung und damit die parallele Skalierbarkeit hängt wesentlich von der Aufteilung der unterschiedlich großen Blöcke auf die einzelnen Prozessoren ab, da deren Behandlung unterschiedlich viele Rechenoperationen und somit verschieden hohe Rechenzeiten erfordern. Daher ist auf dieser mit MPI parallelisierten höchsten Ebene ein Werkzeug zur intelligenten Aufteilung der Blöcke auf die einzelnen MPI-Prozesse implementiert: Liegen beispielsweise ein großer Block und zwei kleinere Blöcke vor, so kann der große Block einem einzelnen Knoten zugeteilt werden, während ein anderer Knoten mit ebenso vielen Kernen die zwei kleinen Blöcke erhält. Benötigen die zwei kleinen Blöcke und der große Block die gleiche Anzahl an Operationen, so ist eine gute Lastverteilung hergestellt.

Da bei großen Parallelrechnern die Anzahl der zur Verfügung stehenden Knoten die Anzahl der Blöcke um ein Vielfaches übersteigt, ist eine Aufspaltung großer Blöcke zusätzlich möglich. Somit können die Verluste in der Skalierbarkeit weiter verringert werden. Hierbei ist jedoch zu beachten, dass die Aufspaltung großer Blöcke aufgrund vermehrter Überlagerungen einen größeren Overhead, also eine zusätzliche zur Prozesskommunikation benötigte Datenbehandlung nach sich zieht.

Eine Ebene tiefer, auf Ebene der Mehrkernprozessoren, ermöglicht nun die Parallelisierung mittels OpenMP eine weitere Verbesserung der Lastverteilung und somit der Skalierbarkeit. Besitzt ein Knoten beispielsweise vier Hauptprozessoren mit je vier Kernen, so ist eine Parallelisierung mit sechzehn OpenMP-Threads sinnvoll. In diesem Fall ist es etwa möglich, dass einem großen Block zwölf Kerne zugeteilt werden, während ein kleiner Block lediglich vier Kerne erhält. Benötigt der kleine Block nur ein Drittel der Operationen des großen, so ist das Lastgleichgewicht hergestellt. Die Kerne kommunizieren hier über den gemeinsamen Speicher des Knoten (*shared memory*) (siehe [46]).

5 Auswertung der implementierten Verfahren

Im Folgenden sollen die im Rahmen dieser Arbeit implementierten und in Kapitel 4 vorgestellten Algorithmen anhand von Testbeispielen ausgewertet werden. Es werden zwei Problemstellungen untersucht.

Im ersten Fall handelt es sich um die Optimierung eines Ersatzmodells, welches auf der Grundlage einer explizit gegebenen Testfunktion $f : [0, 1] \times [0, 1] \rightarrow \mathbb{R}$ erstellt wurde. Diese Funktion wird am Institut für Antriebstechnik des DLR ebenfalls zu Testzwecken verwendet.

Im darauffolgenden Abschnitt 5.2 betrachten wir ein Ersatzmodell basierend auf realen Daten des Verdichterdesigns. Die Stützstellen des Direct Gradient Enhanced Kriging-Verfahrens (Direct GEK; vgl. Abschnitt 3.2) sind somit Verdichterprofile, welche zuvor mit Hilfe eines Strömungslösers berechnet wurden. Die während der Optimierung des Ersatzmodells gewonnenen Resultate werden im Anschluss mit Hilfe des Strömungslösers *MISES* berechnet und bewertet. Hierbei handelt es sich um ein am Massachusetts Institute of Technology (MIT) entwickeltes Programmpaket zur Analyse und Berechnung der Strömung eines Turbinen-Schaufelgitters (siehe [47]).

Zur Charakterisierung des Konvergenzverhaltens der implementierten Verfahren untersuchen wir im Folgenden den Funktionswert des Ersatzmodells als Funktion der Iterationsschritte. Hierbei betrachten wir sowohl die äußeren Iterationen des Augmented-Lagrange-Verfahrens respektive der Barriere-Methode als auch die inneren Iterationen der unrestringierten Verfahren aus Abschnitt 4.3.

5.1 Diskussion anhand einer Testfunktion

Bei der im Folgenden untersuchten Testfunktion handelt es sich um eine Funktion ohne Verbindung zum Verdichterdesign, die zu reinen Testzwecken konzipiert wurde. Die Funktion ist für $(x_1, x_2) \in [0, 1] \times [0, 1]$ wie folgt definiert:

$$f(x_1, x_2) = 18(x_2 - 0.5)^2 - \sqrt{x_1(9(x_2 - 0.5)^2 + 1)} - x_1 \sin(10\pi x_1) + 2.$$

In x_1 -Richtung erhalten wir somit eine Art verzerrte und verschobene Sinusfunktion, in x_2 -Richtung eine Funktion parabelähnlicher Gestalt (siehe Abb. 5.1).

Anhand von dreißig beliebig in der Menge $[0, 1] \times [0, 1] \subseteq \mathbb{R}^2$ gewählten Punkten, den späteren Stützstellen für das Direct GEK, wird nun die Datenbasis zur Erstellung des Ersatzmodells generiert. An diesen Punkten werden somit die Funktionswerte $f(x_1, x_2)$ sowie die partiellen Ableitungen $\frac{\partial f}{\partial x_1}(x_1, x_2)$ und $\frac{\partial f}{\partial x_2}(x_1, x_2)$ berechnet respektive beprobt. Die hier verwendete Datenbasis beziehungsweise die Stützstellen sind in Abbildung 5.2, rechts, durch die schwarzen Punkte gekennzeichnet.

Auf Grundlage dieser an den Stützstellen berechneten Funktionswerte und partiellen Ableitungen wird mit Hilfe des Direct GEK das Ersatzmodell erstellt (vgl. Abschnitt 3.2.5). Hierbei nutzen wir das bereits in AutoOpti implementierte Direct GEK-Programm. Das so erstellte Ersatzmodell ist in Abbildung 5.2, links, dargestellt. Somit ist es nun möglich - ohne Kenntnis der Testfunktion f - an jedem Punkt in $[0, 1] \times [0, 1]$ mit Hilfe des Ersatzmodells eine Approximation des Funktionswertes sowie dessen Unsicherheit zu errechnen. Die Unsicherheit ist ein Maß für die Vertrauenswürdigkeit des Ersatzmodells an jedem Punkt. Sie ist in Abbildung 5.2, rechts, für jeden Punkt aus $[0, 1] \times [0, 1]$ dargestellt und

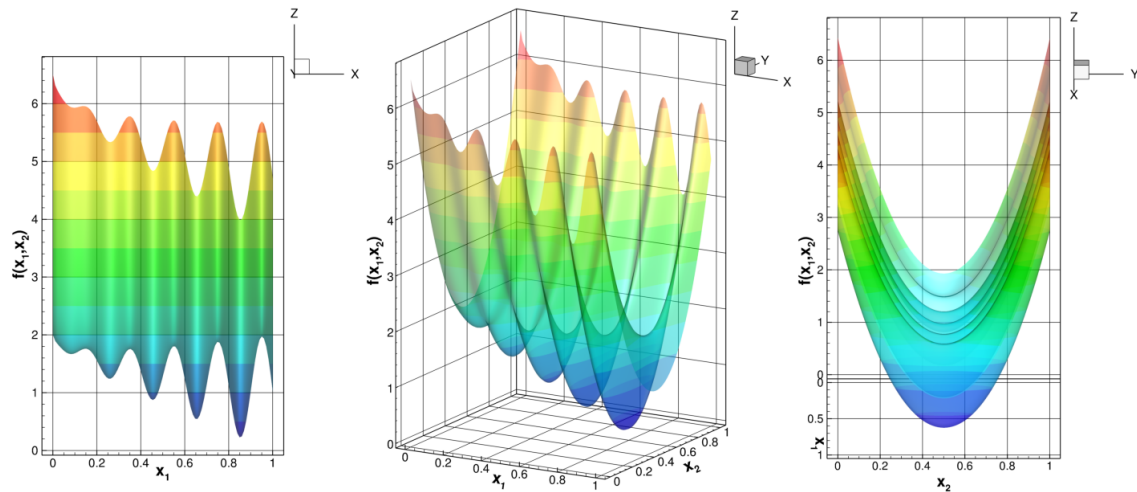


Abbildung 5.1: Graphische Darstellung der Testfunktion in der x_1 - $f(x_1, x_2)$ -Ebene, schräg von der Seite betrachtet und in der x_2 - $f(x_1, x_2)$ -Ebene leicht angeschrägt. Die Farben geben die Höhe der Funktionswerte zusätzlich an.

soll später restringiert werden. Die Grafik zeigt u.a., dass die Unsicherheit nahe eines beprobten Punktes geringer ist als in Gebieten ohne Beprobung. Dies ist intuitiv sinnvoll.

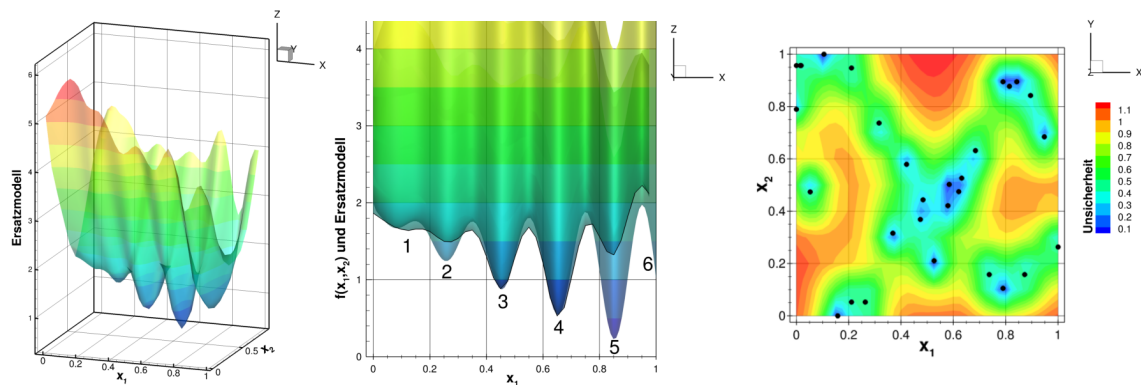


Abbildung 5.2: links: Schrägansicht des Ersatzmodells, Mitte: Vergleich $f(x_1, x_2)$ und Ersatzmodell in der x_1 - $f(x_1, x_2)$ -Ebene, rechts: Datenbasis bzw. Stützstellen und Unsicherheit der Funktionswerte des Ersatzmodells in der x_1 - x_2 -Ebene.

Abbildung 5.2, Mitte, zeigt, dass es sich bei dem Ersatzmodell tatsächlich nur um eine Approximation an f handelt. Wir sehen die vergrößerte x_1 - $f(x_1, x_2)$ -Ebene. Abgetragen werden die Funktion f und das Ersatzmodell, welches mit Hilfe der schwarzen Linie hervorgehoben wird. Bezeichnen wir die Minima von links nach rechts mit den Ziffern 1 bis 6, so stellt das Ersatzmodell für die Minima 3 und 4 eine gute Approximation an die Funktion f dar. Nahe der ersten beiden Minima ist das Ersatzmodell „verwaschen“. Minimum 1 wird nur sehr ungenau und nach rechts verschoben sichtbar, Minimum 2 wurde zwar vom Ersatzmodell richtig erfasst, allerdings mit einem zu hohen Funktionswert versehen. Ähnliches ist bei Minimum 5 geschehen. Das eigentlich noch vorhandene Minimum 6

findet sich im Ersatzmodell nicht wieder. Diese Beobachtungen stimmen tatsächlich auch mit den Unsicherheiten in Abbildung 5.2, rechts, überein. In der Umgebung der Menge $\{(x_1, x_2) \in [0.35, 0.7] \times \{0.5\}\}$, also im Bereich der Minima 3 und 4, befinden sich viele beobachtete Punkte und der Wert der Unsicherheit ist kleiner als 0.6. Entfernen wir uns mit x_1 von diesem Bereich, so vergrößert sich die Unsicherheit.

Im Folgenden betrachten wir unterschiedliche Problemstellungen zur Optimierung des Ersatzmodells der Testfunktion f . Die Analyse und Auswertung dieser Problemstellungen wird zeigen, dass die implementierten Algorithmen das Geforderte, die Minimierung einer nichtlinearen Zielfunktion unter Einhaltung von nichtlinearen Nebenbedingungen, leisten. Das Augenmerk liegt insbesondere im Vergleich der verschiedenen implementierten Verfahren zur restringierten und unrestringierten Minimierung. Hierbei handelt es sich um das Augmented-Lagrange-Verfahren, die Barriere-Methode und die in Abschnitt 4.3 vorgestellten Verfahren.

Zu Beginn schränken wir die Menge der zulässigen Punkte (x_1, x_2) so ein, dass das Ersatzmodell ein eindeutig bestimmtes Minimum besitzt und das Konvergenzverhalten der verschiedenen implementierten Verfahren besser vergleichbar ist. Im Anschluss sind alle Punkte der Menge $[0, 1] \times [0, 1]$ zugelassen und wir beginnen die Optimierung in verschiedenen Startpunkten. Dies zeigt die Unabhängigkeit der Konvergenz vom gewählten Startpunkt. Schließlich geht die Unsicherheit als nichtlineare Nebenbedingung in die Problemstellung ein und das Verhalten bei unterschiedlicher Restriktion dieser Unsicherheit wird untersucht.

Im Folgenden seien die Parameter des Armijo-Verfahrens wie in Abschnitt 4.3.5 angegeben gewählt. Für die restringierten Verfahren findet die prozentuale Abbruchbedingung mit $p_1 = 1$, $p_2 = 0.01$, $p_3 = 1$ und $p_4 = 0.01$ Anwendung (vgl. S. 37). Die Abbruchtoleranzen ω und ω_x der unrestringierten Verfahren setzen wir auf die Werte 10^{-8} respektive $12 \cdot 10^{-9}$ (vgl. S. 40). Diese Parameter gelten, sofern keine anderen Werte angegeben sind, für alle folgenden Unterabschnitte.

5.1.1 Optimierung bei eindeutig bestimmtem Minimum

In diesem Abschnitt untersuchen wir das Verhalten der implementierten Algorithmen bei einer Einschränkung des Untersuchungsgebiets auf die Menge $[0.56, 0.75] \times [0, 1]$. Wie Abbildung 5.3 zeigt, besitzt das Ersatzmodell aufgrund der Einschränkung der freien Variable x auf diesen Bereich ein eindeutig bestimmtes Minimum und das Konvergenzverhalten der verschiedenen Verfahren ist somit besser vergleichbar. Die Unsicherheit soll hier nicht restringiert werden. Der in Abbildung 5.3 schwarz eingezeichnete Punkt $(0.72, 0.86)$ ist Startpunkt aller Verfahren.

Abbildung 5.4, links, zeigt die Konvergenzverläufe des Augmented-Lagrange-Verfahrens bei Verwendung der vier verschiedenen Verfahren zur unrestringierten Minimierung für das zuvor beschriebene Optimierungsproblem. Aus Gründen der Übersichtlichkeit verwenden wir für beide Achsen

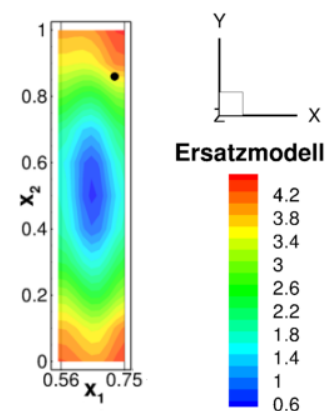


Abbildung 5.3: Niveaumenge des Ersatzmodells.

eine logarithmische Darstellung. Daher kann Iteration 0 nicht abgebildet werden. Hier weisen die vier Verfahren jedoch dasselbe Verhalten auf, da es sich bei allen um den ersten Schritt vom Startpunkt in Richtung des Antigradients handelt (vgl. Verfahren 4.3 bis 4.5). Erst in der sich anschließenden Iteration zeigen die Verfahren ein unterschiedliches Verhalten.

Das Augmented-Lagrange-Verfahren konvergiert mit allen Verfahren zur unrestringierten Minimierung gegen das eindeutige Minimum nahe $(0.655, 0.49)$ mit dem Funktionswert 0.519.

Das Fletcher-Reeves- und das Polak-Ribière-Verfahren zeigen im groben Konvergenzverlauf ein ähnliches Verhalten: Speziell zu Beginn der Optimierung folgt oftmals einer Iteration mit starker Verbesserung des Funktionswertes eine Iteration mit kaum vorhandener Funktionswertverbesserung. Erklären könnte man dieses Verhalten mit Hilfe des CG-Verfahrens für quadratische Zielfunktionen, dessen Verallgemeinerung die beiden Verfahren darstellen (vgl. Abschnitt 4.3.2 und 4.3.3). Die Verfahren von Fletcher-Reeves und Polak-Ribière unterstellen, dass die zu minimierende Funktion speziell nahe dem Minimum eine einer quadratischen Zielfunktion ähnliche Form aufweist. Wie Abbildung 5.5, oben, veranschaulicht, tritt diese quadratische Form insbesondere bei den vom Minimum entfernten Iterierten jedoch nicht auf, sodass hier zwei aufeinanderfolgende Iterierte wiederholt keine wesentliche Funktionswertverbesserung aufweisen. Dieses Verhalten lässt sich besonders während der ersten sechs in Abbildung 5.5, oben rechts, verzeichneten Iterationen des Polak-Ribière-Verfahrens erkennen, aber auch Abbildung 5.5, oben links, zeigt aufeinanderfolgende Fletcher-Reeves-Iterierte ohne Funktionswertverbesserung.

Vergleichen wir die beiden Verfahren untereinander, so zeigt das Polak-Ribière-Verfahren gegenüber dem Fletcher-Reeves-Verfahren eine schnellere Konvergenz mit weniger benötigten Iterationen. Diese langsame Konvergenz des Fletcher-Reeves-Verfahrens lässt sich mit Hilfe des Zick-Zack-Kurses erklären, welchen das Fletcher-Reeves-Verfahren nahe dem Minimum verfolgt (vgl. Abb. 5.5, oben links). Der Grund für dieses Zick-Zack-Verhalten ist nicht eindeutig festzustellen, da das Fletcher-Reeves-Verfahren als Verbesserung des Gra-

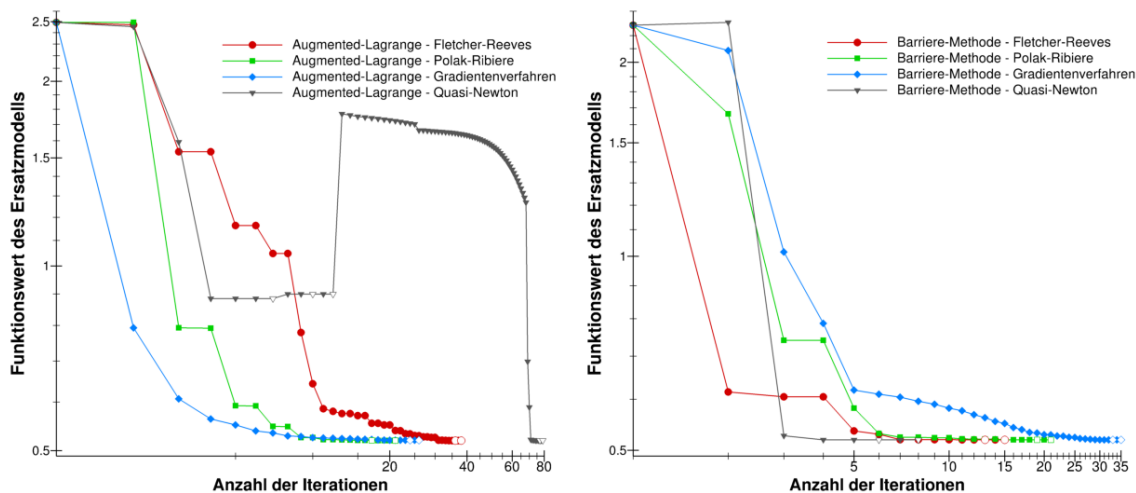


Abbildung 5.4: Konvergenzverhalten des Augmented-Lagrange-Verfahrens (links) und der Barriere-Methode (rechts) bei eindeutig bestimmtem Minimum mit verschiedenen Verfahren zur unrestringierten Minimierung (ausgefülltes Symbol: unrestringiertes Verfahren, nichtausgefülltes Symbol: restringiertes Verfahren).

dientenverfahrens dessen nachteiliges Zick-Zack-Verhalten nahe dem Minimum eigentlich nicht übernehmen sollte. Der dennoch zeitweise in Erscheinung tretende Zick-Zack-Kurs könnte aufgrund der dort fehlenden Ähnlichkeit der zu minimierenden Funktion mit einer quadratischen Zielfunktion auftreten. In diesem Zusammenhang sei darauf hingewiesen, dass das in Abbildung 5.5 verzeichnete Minimum aufgrund einer zu geringen Auflösung des Rechengitters falsch dargestellt wird. Hiervon können wir ausgehen, da alle Verfahren unabhängig voneinander gegen denselben Punkt konvergieren.

Betrachten wir das Gradientenverfahren, so zeigt dieses ein erstaunlich gutes Verhalten. Die Iterationsanzahl ist nahe der des Polak-Ribière-Verfahrens und insbesondere erheblich geringer als die des Fletcher-Reeves-Verfahrens. Dessen ungeachtet zeigt sich hier jedoch schon die langsame Konvergenz des Gradientenverfahrens nahe dem Minimum. Diese lässt sich mit dem bekannten Zick-Zack-Verhalten des Gradientenverfahrens nahe dem Minimum erklären (vgl. Abb. 5.5, links unten).

Da das Augmented-Lagrange-Verfahren auch ein Verlassen des zulässigen Bereichs gestat-

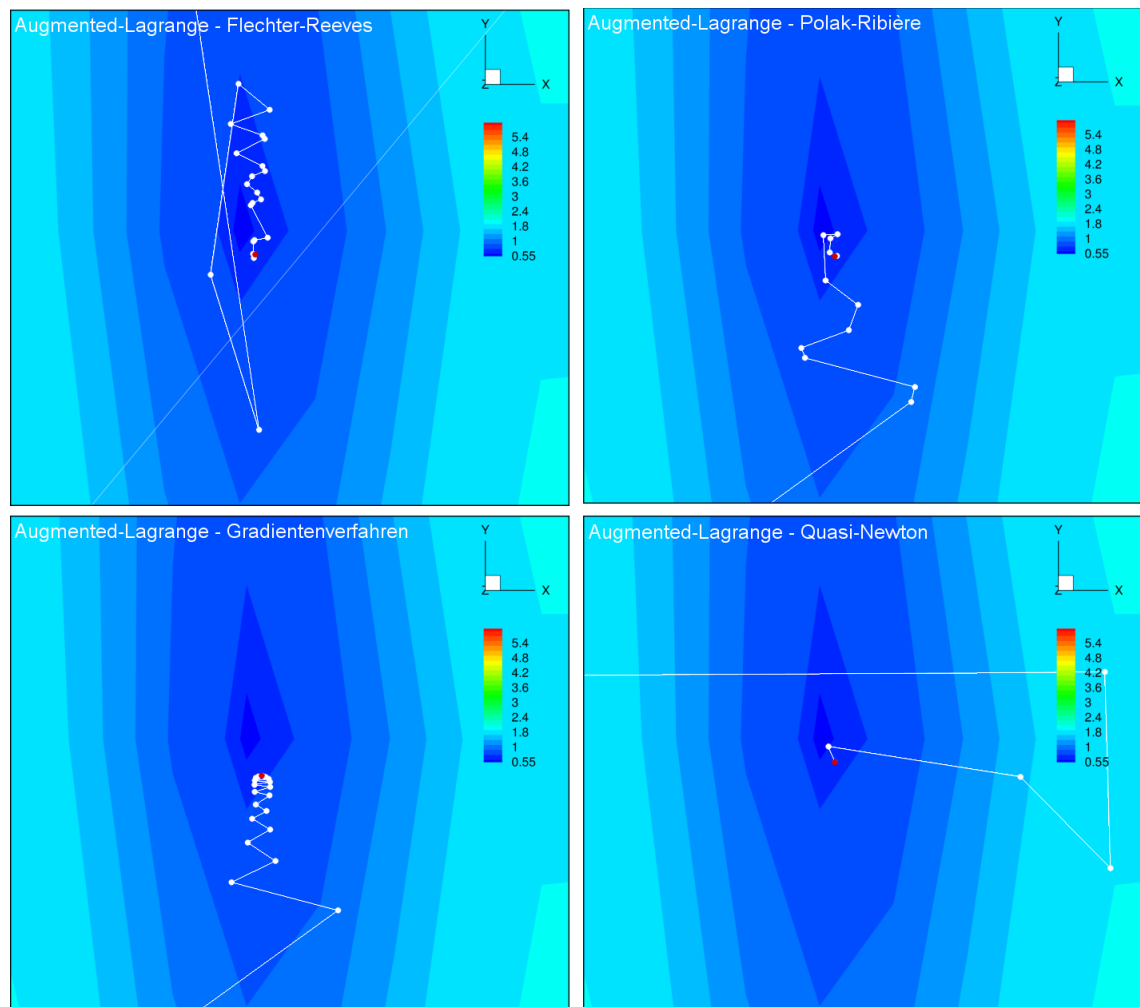


Abbildung 5.5: Vergrößerter Ausschnitt der Niveaulfläche des Ersatzmodells um das Minimum $(0.655, 0.49)$; verzeichnet sind die Augmented-Lagrange-Iterierten aller unrestringierten Verfahren nahe dem rot eingezeichnetem Minimum.

tet, zeigt das Quasi-Newton-Verfahren das in Abbildung 5.4, links, dargestellte Verhalten. Bis einschließlich Iteration 12 befindet sich die aktuelle Iterierte $x = (0.46, 0.52)$ in der Nähe des dritten Minimums (vgl. Abb. 5.2) und somit weit außerhalb des zulässigen Bereichs. Erst eine Erhöhung des Strafparameters μ_k aufgrund der starken Normverletzung (vgl. Schritt 5 in Verfahren 4.2) garantiert letztendlich die Zulässigkeit.

Die darauffolgende langsame Konvergenz lässt sich mit einer zu kleinen Startschrittweite erklären. Abbildung 5.6 zeigt eine signifikante Verbesserung des Konvergenzverhaltens bei einer Vergrößerung von SIGMA_SHORT auf den Wert 0.06 (vgl. Abschnitt 4.3.5, S. 46). Dies führt sodann zu dem in Abbildung 5.5, unten rechts, dargestellten schnellen Konvergenzverhalten. Die Anpassung des Wertes für SIGMA_SHORT kann vom Benutzer über die Eingabedatei auch *online*, d.h. während der laufenden Optimierung, durchgeführt werden. Beobachtet dieser also zu kleine Startschrittweiten, so lässt sich dieses Defizit noch während der Optimierung beheben.

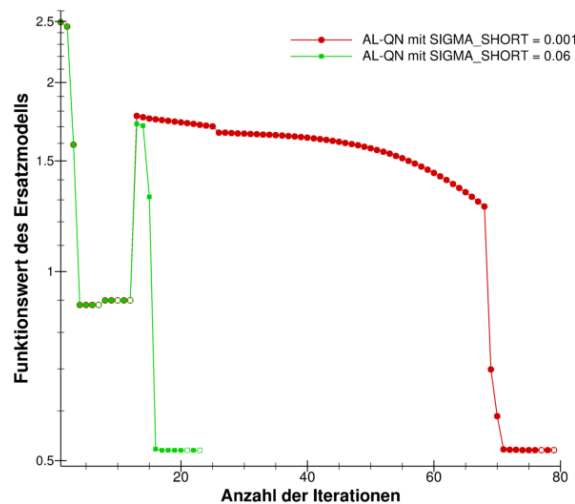


Abbildung 5.6: Konvergenzverläufe des Augmented-Lagrange-Verfahrens mit dem Quasi-Newton-Verfahren zur unrestringierten Minimierung bei verschiedenen Werten für SIGMA_SHORT (ausgefülltes Symbol: Quasi-Newton-Iteration, nichtausgefülltes Symbol: Augmented-Lagrange-Iteration).

Abbildung 5.4, rechts, zeigt die Konvergenzverläufe der Barriere-Methode bei Verwendung der vier verschiedenen Verfahren zur unrestringierten Minimierung. Auch hier wurde eine logarithmische Darstellung beider Achsen gewählt. Alle vier Verfahren konvergieren gegen das Minimum nahe $x = (0.655, 0.49)$ mit dem Funktionswert 0.519.

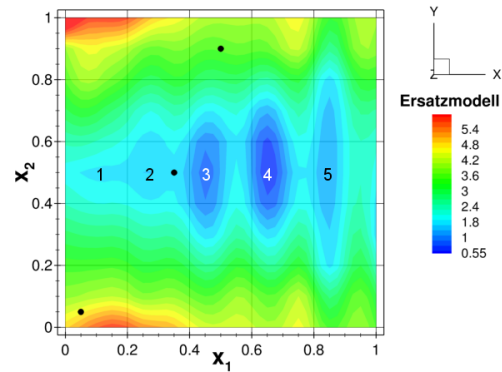
Das Fletcher-Reeves-Verfahren zeigt hier gegenüber dem Polak-Ribière-Verfahren ein besseres Verhalten. Wir erhalten schnelle Konvergenz bei einer geringeren Anzahl an Iterationen. Das Gradientenverfahren weist ein langsames Konvergenzverhalten auf, welches sich insbesondere zwischen Iteration 5 und Iteration 33 zeigt. Das Verfahren genügt der Abbruchbedingung erst nach extrem vielen Iterationen. Das Quasi-Newton-Verfahren zeigt im Vergleich zu den übrigen drei Verfahren das beste Konvergenzverhalten. Schon ab Iteration 3 befindet sich das Verfahren nahe dem Minimum. Insgesamt benötigt es nur 8 Iterationen zur Konvergenz.

Vergleichen wir das Augmented-Lagrange-Verfahren und die Barriere-Methode in Abbildung 5.4, so zeigen die unrestringierten Verfahren bei Verwendung der Barriere-Methode

grundsätzlich ein besseres Verhalten als beim Augmented-Lagrange-Verfahren, d.h. schnellere Konvergenz und durchschnittlich eine geringere Anzahl an Iterationen. Unter den unrestringierten Verfahren weist das Quasi-Newton-Verfahren das beste Konvergenzverhalten auf (beim Augmented-Lagrange-Verfahren nach Anpassung von SIGMA_SHORT).

5.1.2 Optimierung bei verschiedenen Startpunkten

In diesem Abschnitt untersuchen wir das Verhalten der verschiedenen Algorithmen für unterschiedliche Startpunkte ohne weitere Restriktion an den Definitionsbereich und die Unsicherheit, d.h. $x \in [0, 1] \times [0, 1]$. Für diesen Fall zeigt Abbildung 5.7 die Niveaumenge des Ersatzmodells und die drei analysierten Startpunkte $(0.5, 0.9)$, $(0.05, 0.05)$ und $(0.35, 0.5)$.



Startpunkt $(0.5, 0.9)$

Wir beginnen mit der Untersuchung des Startpunktes $(0.5, 0.9)$. Abbildung 5.8 zeigt die Konvergenzverläufe des Augmented-Lagrange-Verfahrens und der Barriere-Methode für alle vier Verfahren zur unrestringierten Minimierung. Vorweg fällt auf, dass die unterschiedlichen Optimierungsverfahren gegen drei verschiedene Minima konvergieren. Hierbei handelt es sich um Minimum 3 nahe $x = (0.455, 0.500)$ mit dem Funktionswert 0.88, Minimum 4 bei $x = (0.655, 0.490)$ mit dem Funktionswert 0.52 und Minimum 5 bei $x = (0.840, 0.465)$ mit

Abbildung 5.7: Niveaumenge des Ersatzmodells für $x \in [0, 1] \times [0, 1]$ mit Startpunkten.

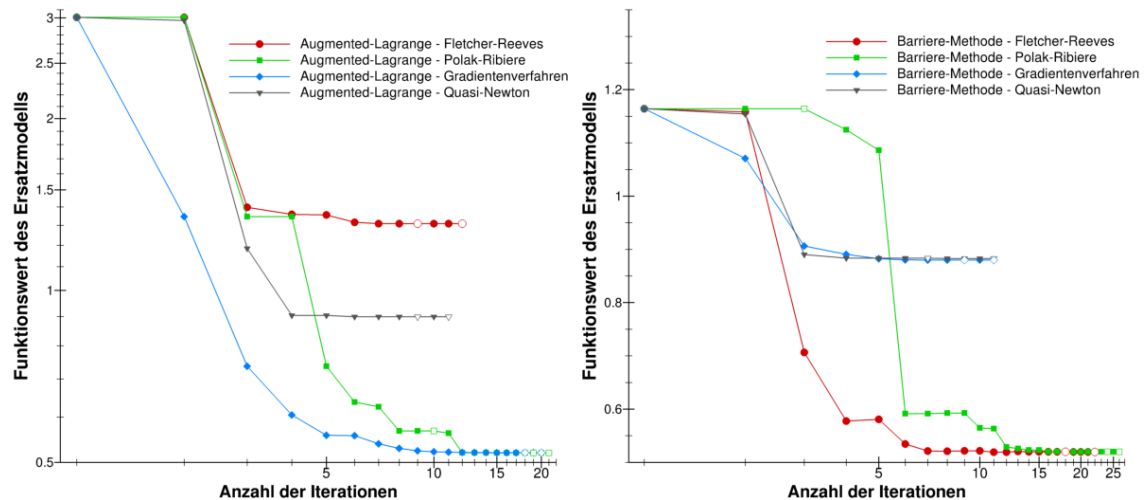


Abbildung 5.8: Konvergenzverläufe des Augmented-Lagrange-Verfahrens (links) und der Barriere-Methode (rechts) mit Startpunkt $(0.5, 0.9)$ und verschiedenen Verfahren zur unrestringierten Minimierung (ausgefülltes Symbol: unrestringiertes Verfahren, nichtausgefülltes Symbol: restringiertes Verfahren).

dem Funktionswert 1.31. Dies ist nicht verwunderlich, da die implementierten Verfahren gegen lokale und nicht globale Minima konvergieren. Das Zustandekommen unterschiedlicher Minima ist auf die verschiedenen Abstiegsrichtungen d der einzelnen Verfahren zur unrestringierten Minimierung und die davon abhängigen und somit differierenden Startschrittweiten zurückzuführen.

Grundsätzlich zeigen alle Verfahren ein gutes Konvergenzverhalten mit einer geringen Anzahl an Iterationen. Insbesondere das Quasi-Newton-Verfahren, welches schon im vorherigen Abschnitt nach Änderung von SIGMA_SHORT ein vergleichbar gutes Konvergenzverhalten zeigte, besticht aufgrund der bei beiden Verfahren geringen Anzahl an Iterationen (beim Augmented-Lagrange-Verfahren nach Änderung von SIGMA_SHORT auf den Wert 0.03).

Startpunkte (0.05,0.05) und (0.35,0.5)

Betrachten wir nun die beiden Startpunkte (0.05, 0.05) bzw. (0.35, 0.5) (Konvergenzverläufe siehe Abb. 5.9, oben respektive unten), so zeigen die Verfahren auch hier ein gutes Konvergenzverhalten. Bemerkenswert ist die durchschnittlich sehr gute Konvergenz des Quasi-Newton-Verfahrens für alle vier Problemstellungen (für den Startpunkt (0.05, 0.05) ist beim Augmented-Lagrange-Verfahren aufgrund zu kleiner Startschrittweiten eine Vergrößerung von SIGMA_SHORT auf 0.005 nötig). Das Quasi-Newton-Verfahren weist in allen Fällen die geringste Anzahl an Iterationen auf und vermittelt einen stabilen Eindruck.

Das Gradientenverfahren konvergiert zeitweise sehr gut, insbesondere für den Startpunkt (0.05, 0.05) beim Augmented-Lagrange-Verfahren (vgl. Abb. 5.9, oben links). Insgesamt zeigt sich allerdings nahe dem Minimum eine langsame Konvergenz des Verfahrens. Im Vergleich zu den beiden verbleibenden Verfahren von Fletcher-Reeves und Polak-Ribière, welche eine durchschnittlich hohe Anzahl an Iterationen aufweisen, scheint es sich beim Gradientenverfahren jedoch um die stabilere Methode zu handeln.

Das Verhalten des Polak-Ribière-Verfahrens zum Startpunkt (0.05, 0.05) beim Augmented-Lagrange-Verfahren (vgl. Abb. 5.9, links oben) kann wiederum auf eine zeitweise Unzulässigkeit der Iterierten zurückgeführt werden. Bis zur Iteration 15 befindet sich das Verfahren in der Nähe des Punktes (1.06, 0.30) und somit außerhalb der zulässigen Menge $[0, 1] \times [0, 1]$. Erst eine Erhöhung des Strafparameters μ_k hat eine Zulässigkeit der Iterierten aber somit auch einen erhöhten Funktionswert zur Folge. Das Verfahren konvergiert letztendlich gegen das Minimum 1 bei (0.16, 0.49) mit dem Funktionswert 1.66.

Zusammenfassend folgt aus dem vorausgehenden Abschnitt eine Unabhängigkeit des groben Konvergenzverhaltens vom Startpunkt. Im Detail zeigen sich natürlich Unterschiede der verschiedenen Verfahren, allerdings beobachten wir bei allen Verfahren Konvergenz in einer annehmbaren Anzahl an Iterationen. Bei den unrestringierten Verfahren ist insbesondere das Quasi-Newton-Verfahren positiv hervorzuheben, welches durchweg eine geringe Iterationsanzahl und ein stabiles Konvergenzverhalten aufweist.

Betrachten wir die Barriere-Methode und das Augmented-Lagrange-Verfahren im Vergleich, so zeigt die zuerst genannte Methode durchschnittlich ein besseres oder zumindest gleich gutes Verhalten. Dies konnten wir schon in Abschnitt 5.1.1 beobachten.

Im Folgenden werden wir untersuchen, inwieweit eine Restriktion der Unsicherheit dieses Verhalten beeinträchtigt.

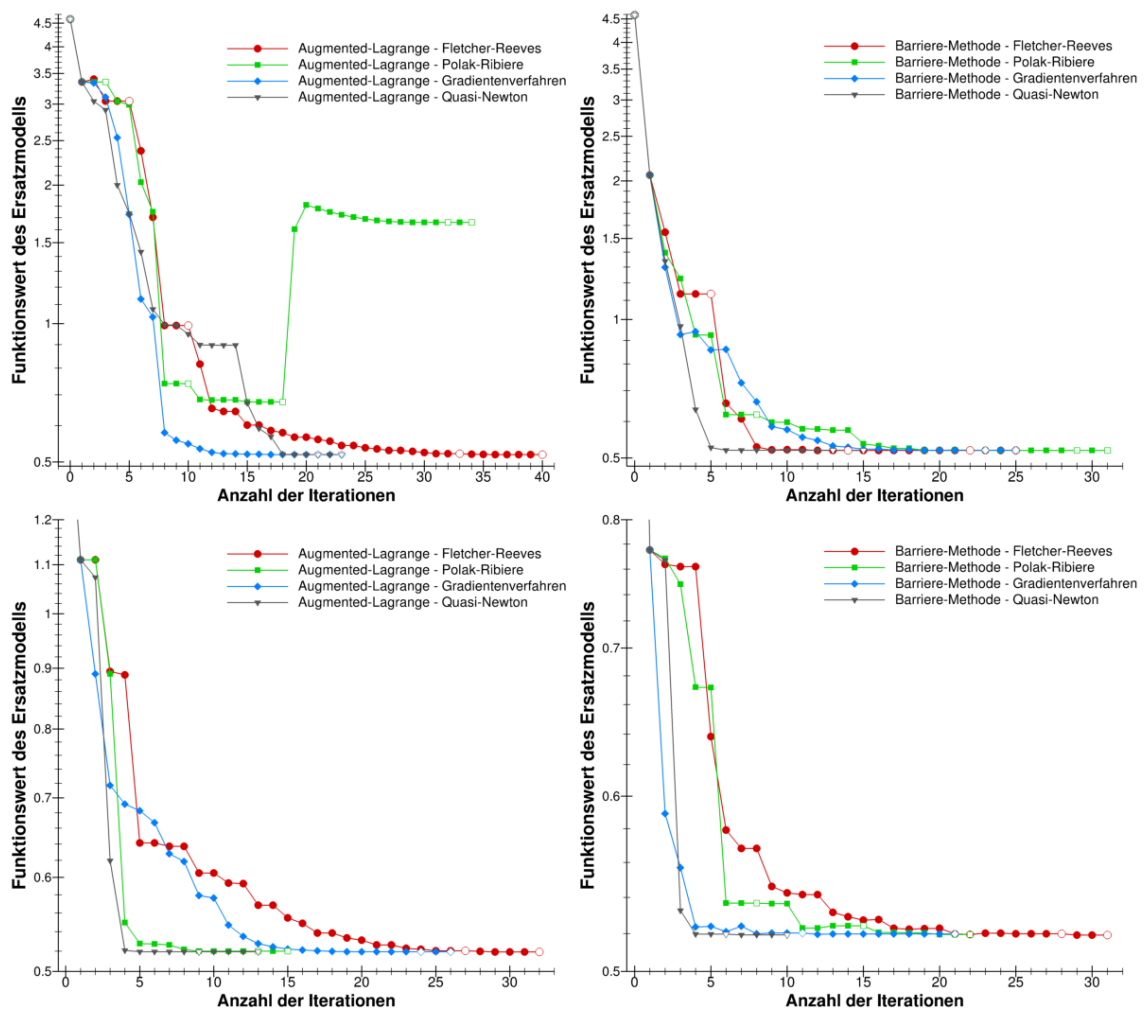


Abbildung 5.9: Konvergenzverläufe des Augmented-Lagrange-Verfahrens (jeweils links) und der Barriere-Methode (jeweils rechts) mit Startpunkt $(0.05, 0.05)$ (jeweils oben) respektive $(0.35, 0.5)$ (jeweils unten) und verschiedenen Verfahren zur unrestringierten Minimierung (ausgefülltes Symbol: unrestringiertes Verfahren, nichtausgefülltes Symbol: restringiertes Verfahren).

5.1.3 Optimierung bei Restriktion der Unsicherheit

Aufgrund der Optimierung von Ersatzmodellen, welche eine Approximation an die wahre Funktion darstellen, sind die gewonnenen Ergebnisse immer mit einer Unsicherheit versehen. Diese Unsicherheit, in unserem Beispiel eine reelle Zahl zwischen 0 und 1.15, erhalten wir als zusätzliche Information aus dem Direct GEK-Verfahren. Das Ziel ist nun eine Minimierung des Funktionswertes des Ersatzmodells, wobei die Unsicherheit einen gewissen Wert nicht übersteigen soll.

Im Folgenden untersuchen wir das Verhalten der verschiedenen Algorithmen bei unterschiedlichen Werten für die Restriktion der Unsicherheit.

Restriktion der Unsicherheit durch 0.7

Zunächst begrenzen wir die Unsicherheit durch den Wert 0.7 nach oben, d.h. ist die Unsicherheit durch die Funktion $h : [0, 1] \times [0, 1] \rightarrow \mathbb{R}$ gegeben, so soll

$$h(x) \leq 0.7$$

für alle $x \in [0, 1] \times [0, 1]$ gelten. Abbildung 5.10, links, gibt die Unsicherheiten für alle Punkte aus $[0, 1] \times [0, 1]$ wieder. Bei den dunkelroten Flächen handelt es sich um Punkte, die die Nebenbedingungsrestriktion nicht erfüllen, d.h. für diese Punkte gilt $h(x) > 0.7$. Wir wählen als Startpunkt den Punkt $(0.32, 0.7)$, der in Abbildung 5.10 schwarz eingezeichnet ist.

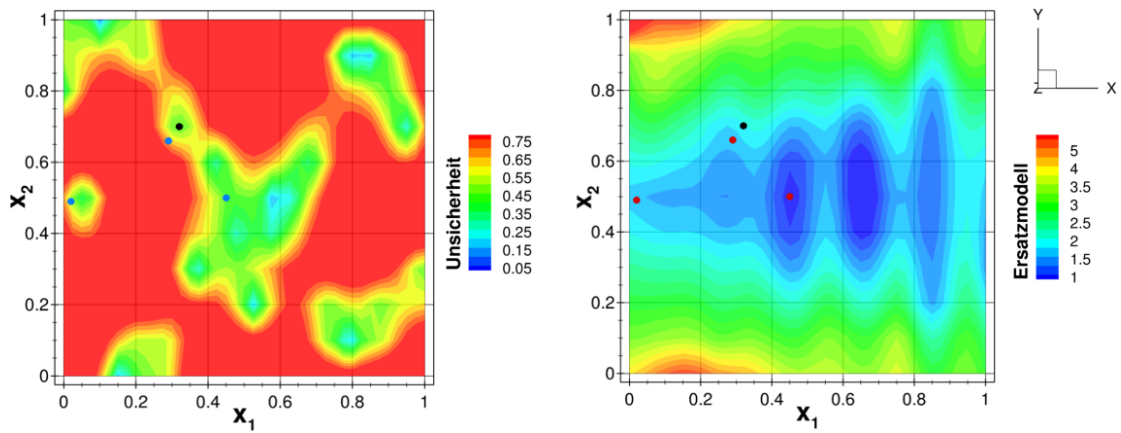


Abbildung 5.10: Niveauflächen der Unsicherheit (links) und des Ersatzmodells (rechts) bei einer Unsicherheitsrestriktion von 0.7 mit schwarz eingezeichnetem Startpunkt $(0.32, 0.7)$ und den drei erreichten Zielpunkten (eingezeichnet in blau respektive rot).

Abbildung 5.11, links, zeigt die Konvergenzverläufe des Augmented-Lagrange-Verfahrens für die beschriebene Problemstellung.

Das Fletcher-Reeves-Verfahren konvergiert gegen den in Abbildung 5.10, links, blau eingezeichneten Punkt $(0.29, 0.66)$. Betrachten wir den Konvergenzverlauf, so zeigt dieser das typische Verhalten aufgrund einer Unzulässigkeit der Iterierten. Wie der schwarz eingezeichnete Verlauf der Fletcher-Reeves-Iterierten in Abbildung 5.12, links, zeigt, befindet sich das Verfahren tatsächlich vorerst in der Nähe des Punktes $(0.12, 0.50)$. Dieser verletzt mit einer zu hohen Unsicherheit von 0.74 die Nebenbedingung. Erst eine zweifache Erhöhung des Strafparameters μ_k führt zu einer Zulässigkeit der Iterierten. Das Verfahren springt daraufhin zu dem Punkt $(0.30, 0.81)$ (vgl. Abb. 5.12) und konvergiert letztendlich gegen den Punkt $(0.29, 0.66)$ mit dem Funktionswert 1.8.

Den Grund für dieses Konvergenzverhalten erkennen wir in Abbildung 5.12, rechts. Die Kontur der Niveaumengen des Ersatzmodells und die daraus folgenden Abstiegsrichtungen führen das Verfahren vom Punkt $(0.30, 0.81)$ zu dem Punkt $(0.29, 0.66)$. Eine weitere Verkleinerung des Funktionswertes wird an diesem Punkt jedoch durch die Restriktion der Unsicherheit (vgl. Abb. 5.12, links) verhindert und das Verfahren konvergiert gegen ein lokales Minimum.

Betrachten wir für das Augmented-Lagrange-Verfahren die Konvergenzverläufe des Polak-Ribière- und des Gradientenverfahrens, so weisen auch diese auf eine vorübergehende Un-

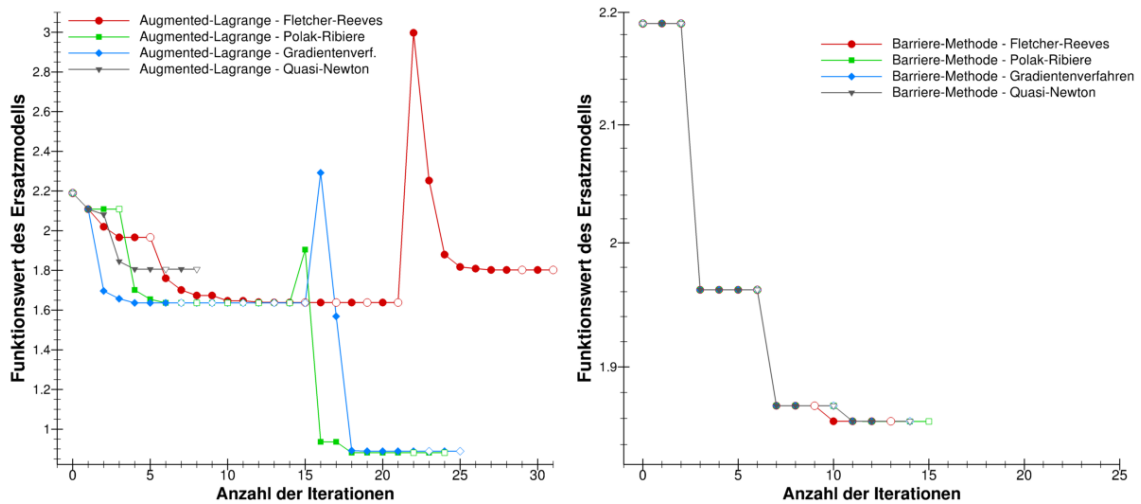


Abbildung 5.11: Konvergenzverläufe des Augmented-Lagrange-Verfahrens (links) und der Barriere-Methode (rechts) bei Restriktion der Unsicherheit durch 0.7 mit verschiedenen Verfahren zur unrestringierten Minimierung (ausgefülltes Symbol: unrestringiertes Verfahren, nichtausgefülltes Symbol: restringiertes Verfahren).

zulässigkeit der Iterierten hin. Beide Verfahren konvergieren nach Erhöhung des Strafparsameters gegen das lokale Minimum bei $x = (0.45, 0.50)$ mit dem Funktionswert 0.88. In Abbildung 5.12 ist mittels der weißen Quadrate exemplarisch der Weg des Polak-Ribière-Verfahrens verzeichnet. Da beide Verfahren grundsätzlich das gleiche Verhalten zeigen, ist dieser Weg auch stellvertretend für das Gradientenverfahren. Beiden Verfahren gelingt es aufgrund eines sich vom Fletcher-Reeves-Verfahren leicht unterscheidenden Ausgangs-

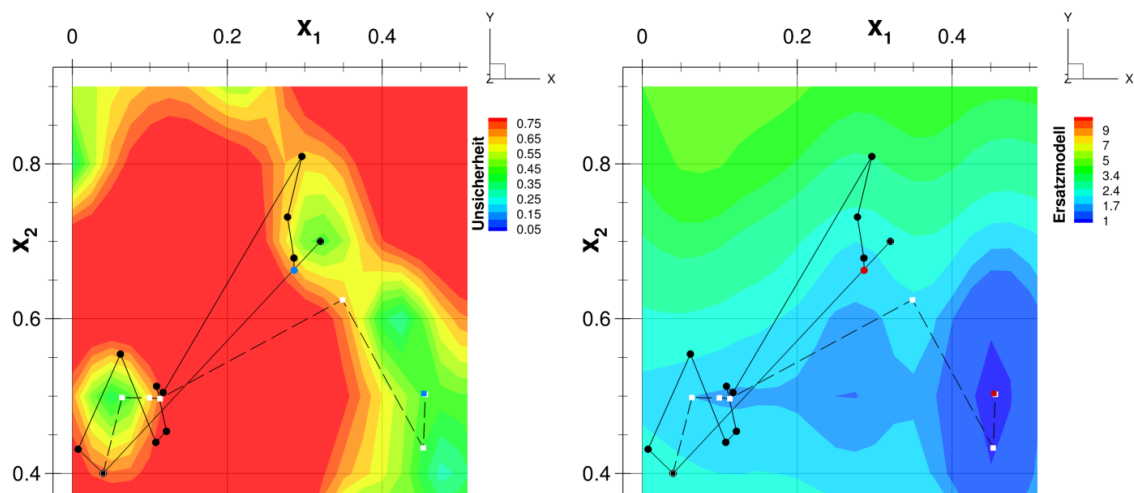


Abbildung 5.12: Vergrößerter Ausschnitt des Untersuchungsgebietes mit Verlauf der Iterierten beim Augmented-Lagrange-Verfahren mit Fletcher-Reeves-Verfahren (Weg mit schwarzen Punkten) und Polak-Ribière-Verfahren (gestrichelter Weg mit weißen Quadraten), links: Niveauebene der Unsicherheit, rechts: Niveauebene des Ersatzmodells.

punktes (0.11, 0.49) die Senke der Unsicherheit in Richtung des Punktes (0.35, 0.62) zu verlassen (vgl. Abb. 5.12, links). Infolge der Kontur der Niveaumengen des Ersatzmodells in diesem Punkt konvergieren beide Verfahren nun gegen das lokale Minimum bei $x = (0.45, 0.50)$, ohne weitere Restriktionen der Unsicherheit zu verletzen (vgl. Abb. 5.12, rechts).

Die Lage des in Abbildung 5.10, links, blau eingezeichneten Endpunktes (0.02, 0.49) des Augmented-Lagrange-Verfahrens mit Quasi-Newton ist zunächst verwunderlich, insbesondere da Abbildung 5.10, rechts, einen in die Menge $[0, 1] \times [0, 1]$ zeigenden Antigradienten suggeriert. Wie wir jedoch bei genauer Betrachtung der einzelnen Werte feststellen, ist der Gradient des Ersatzmodells im Endpunkt (0.02, 0.49) des Quasi-Newton-Verfahrens der Vektor $(0.36, 0.08)^T$. Folglich zeigt eine Abstiegsrichtung aus der zulässigen Menge. Eine weitere Minimierung in dieser Richtung wird aufgrund der Restriktion der Unsicherheit verhindert (vgl. Abb. 5.10, links). Somit konvergiert das Quasi-Newton-Verfahren gegen das lokale Minimum (0.02, 0.49) mit dem Funktionswert 1.8.

Betrachten wir die Konvergenzverläufe der Barriere-Methode in Abbildung 5.11, rechts, so erkennen wir, dass alle Verfahren gegen den gleichen Punkt mit Funktionswert 1.8 konvergieren. Hierbei handelt es sich um den in Abbildung 5.10, links, blau eingezeichneten Punkt (0.29, 0.66). Dies erklärt sich aus der strikten Zulässigkeit der Barriere-Methode und der vom Startpunkt (0.32, 0.70) ausgehenden Abstiegsrichtung (vgl. Abb. 5.10, rechts). Die einzelnen Verfahren zeigen hier alle ein sehr gutes und während der ersten Iterationen gleiches Konvergenzverhalten, wobei das Fletcher-Reeves-Verfahren die geringste Anzahl an Iterationen aufweist.

Durchschnittlich weist die Barriere-Methode im Vergleich zum Augmented-Lagrange-Verfahren eine geringere Iterationsanzahl auf. Einzig das Quasi-Newton-Verfahren konvergiert bei letztgenannter Methode schneller. Das gute Konvergenzverhalten der Barriere-Methode könnte darin begründet liegen, dass die Iterierten direkt gegen die durch die Unsicherheitsrestriktion auftretende Barriere anlaufen und der Straffparameter r nur noch kontinuierlich verkleinert werden muss, bis das Abbruchkriterium erreicht ist. Ein Nachteil der Barriere-Methode ist der Umstand, dass diese es nicht schafft, sich aus der Umgebung des Startpunktes in Richtung des lokalen Minimums (0.45, 0.50) zu entfernen. Hier ist die Schwelle um den Punkt (0.36, 0.66) aufgrund der hohen Unsicherheit von 0.65 anscheinend zu groß, als dass die logarithmische Barriere-Methode diese überwinden kann (vgl. Abb. 5.10, rechts).

Vergleichen wir die unrestringierten Verfahren, so unterscheidet sich die Iterationsanzahl des Fletcher-Reeves- und des Quasi-Newton-Verfahrens bei der Barriere-Methode nur um eine Iteration. Beim Augmented-Lagrange-Verfahren konvergiert das Quasi-Newton-Verfahren hingegen am schnellsten.

Restriktion der Unsicherheit durch 0.8

Im Folgenden soll das Verhalten der Algorithmen bei Vergrößerung der Unsicherheitsrestriktion auf den Wert 0.8 untersucht werden. Die daraufhin kleinere Menge der unzulässigen Punkte $\{x \in [0, 1] \times [0, 1] \mid h(x) > 0.8\}$ ist in Abbildung 5.13, links, in dunkelrot gekennzeichnet. Der Startpunkt (0.32, 0.70) aller Verfahren ist wiederum in schwarz dargestellt, alle vier durch die verschiedenen Verfahren erreichten Endpunkte sind in Abbildung 5.14 in blau respektive rot eingezeichnet.

Das Augmented-Lagrange-Verfahren mit dem Fletcher-Reeves-Verfahren zur unrestringierten Minimierung ist das einzige Verfahren, welches gegen das lokale Minimum bei $(0.655, 0.490)$ konvergiert. Der Weg der Iterierten ist in Abbildung 5.13 gekennzeichnet. Wir erkennen, dass das Verfahren aufgrund der Schrittweitenberechnung relativ starke Sprünge ausführt. Die zeitweise Unzulässigkeit der Iterierten aufgrund des Verlassens der zulässigen Menge $[0, 1] \times [0, 1]$ am rechten Rand führt zu dem in Abbildung 5.14, links, dargestellten Konvergenzverhalten. Eine Vergrößerung des Strafparameters in Iteration 8 erhöht den Funktionswert vorübergehend. Im Anschluss verkleinert sich dieser jedoch bis zur Konvergenz. Negativ anzumerken sei in diesem Zusammenhang das langsame Konvergenzverhalten nahe dem Minimum ab Iteration 11.

Die drei verbleibenden Verfahren konvergieren gegen das in Abbildung 5.13, links, blau eingezeichnete lokale Minimum bei $(0.11, 0.49)$ mit dem Funktionswert 1.63 (vgl. Abb. 5.14, links). Nach der Ausführung des für alle Verfahren gleichen ersten Schritts in Richtung des Antigradients finden diese aufgrund der vom Fletcher-Reeves-Verfahren differierenden Abstiegsrichtungen nicht aus der Senke der Unsicherheit (vgl. Abb. 5.13, links) und konvergieren daraufhin gegen den angegebenen Punkt. Ein weiterer Abstieg in Richtung größerer x_1 -Werte wird aufgrund der Restriktion der Unsicherheit verhindert.

Die Verfahren konvergieren alle nach wenigen Iterationen. Bemerkenswert ist die erstaunlich schnelle Konvergenz des Gradientenverfahrens. Beim Quasi-Newton-Verfahren muss SIGMA_SHORT aufgrund zu kleiner Startschrittweiten auf den Wert 0.02 gesetzt werden.

Betrachten wir die Konvergenzverläufe der Barriere-Methode in Abbildung 5.14, rechts, so zeigen die drei ersten Verfahren ein sehr ähnliches Verhalten. Das Fletcher-Reeves-, das Polak-Ribière- und das Gradientenverfahren konvergieren alle gegen den in Abbildung 5.13, links, blau eingezeichneten Punkt $(0.29, 0.64)$. Auch die Anzahl der benötigten Iterationen unterscheidet sich nur jeweils um eine Iteration und ist daher vergleichbar gut. Einzig das Quasi-Newton-Verfahren konvergiert nach dem ersten Schritt in Richtung des Antigradients aufgrund einer daraufhin anderen Abstiegsrichtung gegen das lokale Minimum bei $(0.45, 0.50)$ mit dem Funktionswert 0.88 (vgl. Abb. 5.14, rechts). Der Nachteil des Quasi-Newton-Verfahrens ist die nach Iteration 2 langsame Konvergenz nahe dem Mini-

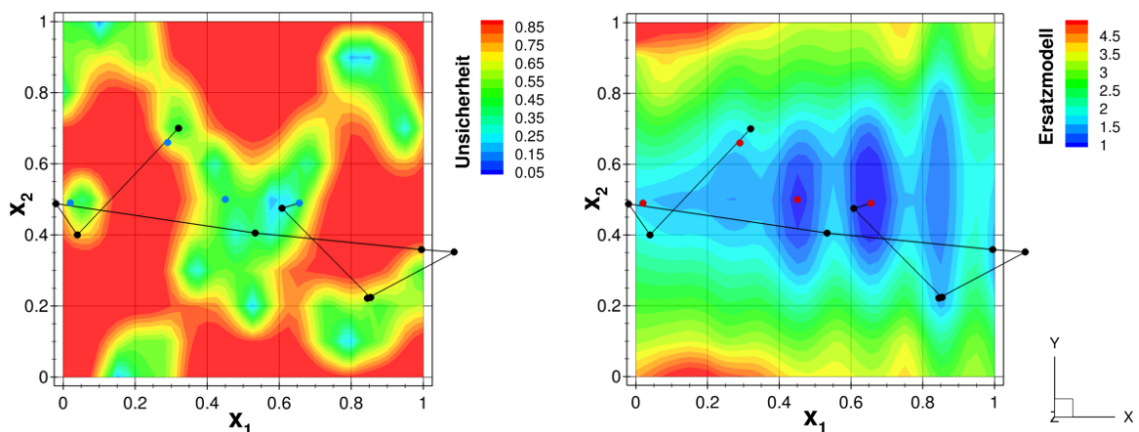


Abbildung 5.13: Verlauf der Iterierten des Augmented-Lagrange-Verfahren mit Fletcher-Reeves-Verfahren (Weg mit schwarzen Punkten), links: Niveauläche der Unsicherheit bei Restriktion durch 0.8, rechts: Niveauläche des Ersatzmodells.

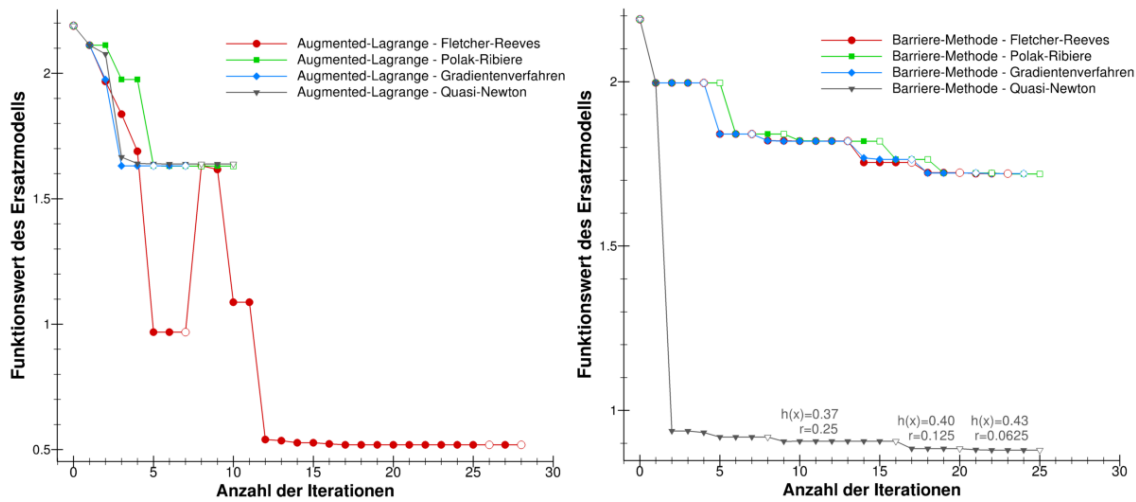


Abbildung 5.14: Konvergenzverläufe des Augmented-Lagrange-Verfahrens (links) und der Barriere-Methode (rechts) bei Restriktion der Unsicherheit durch 0.8 mit verschiedenen Verfahren zur unrestringierten Minimierung (ausgefülltes Symbol: unrestringiertes Verfahren, nichtausgefülltes Symbol: restringiertes Verfahren).

mum. Diese ist jedoch auf die Barriere-Methode und nicht auf das Quasi-Newton-Verfahren zurückzuführen: Bei den Iterationen 8, 16 und 20 handelt es sich jeweils um eine Barriere-Iteration, in welcher der Straffparameter r auf 0.25, 0.125 respektive 0.0625 herabgesetzt wird (vgl. Abb. 5.14, rechts, und Verfahren 4.1). Dies hat eine immer geringere Bedeutung des logarithmischen Barriere-Terms in der Barriere-Funktion Φ zur Folge. Daher konvergiert das Quasi-Newton-Verfahren, welches in jeder Barriere-Iteration die Funktion Φ minimiert, gegen Punkte, deren Unsicherheit sich von Barriere-Iteration zu Barriere-Iteration erhöht. Starten wir nach Iteration 8 mit einer Unsicherheit von $h(x) = 0.37$, so erhöht sich diese nach Iteration 16 auf $h(x) = 0.40$ und nach Iteration 20 auf $h(x) = 0.43$. Einzeln betrachtet konvergiert das Quasi-Newton-Verfahren zwischen den Barriere-Iterationen in wenigen Iterationen.

Erstmalig zeigt das Augmented-Lagrange-Verfahren im Vergleich zur Barriere-Methode ein besseres Konvergenzverhalten mit durchschnittlich weniger Iterationen. Das Gradientenverfahren konvergiert erstaunlich schnell, das Quasi-Newton-Verfahren mit drei zusätzlichen Iterationen in einer annehmbaren Anzahl an Iterationen. Bei der Barriere-Methode benötigt das Quasi-Newton-Verfahren ähnlich viele Iterationen wie die anderen unrestringierten Verfahren. Es ist das einzige Verfahren, welches sich aus der Umgebung des Startpunktes zu entfernen vermag.

Restriktion der Unsicherheit durch 1.0

In diesem Abschnitt vergrößern wir die Restriktion der Unsicherheit, sodass für alle zulässigen Punkte $h(x) \leq 1.0$ gelten soll. Wie Abbildung 5.15, links, veranschaulicht, verkleinert sich die Menge der in dunkelrot gekennzeichneten unzulässigen Punkte weiter. Der Startpunkt aller Verfahren ist wieder der in Abbildung 5.15 schwarz eingezeichnete Punkt $(0.32, 0.70)$. Des Weiteren sind in der Abbildung alle Endpunkte der verschiedenen

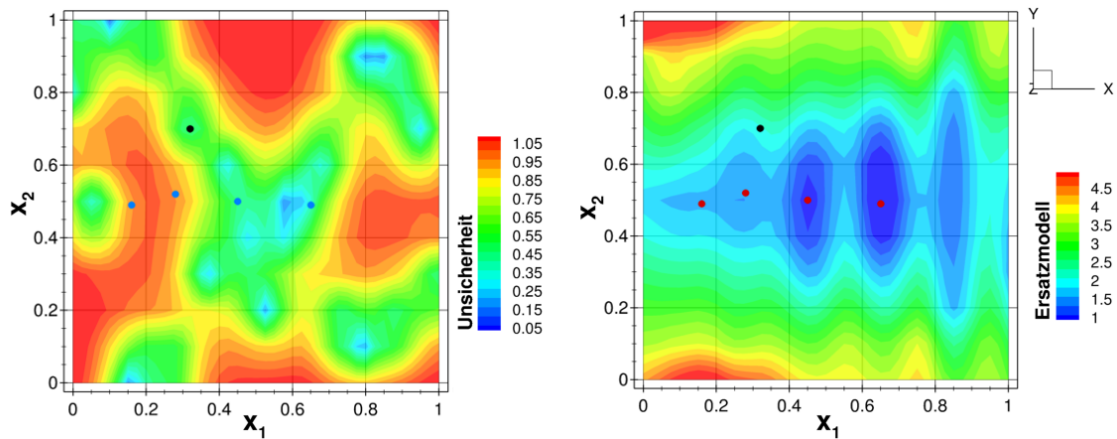


Abbildung 5.15: Niveaüfläche der Unsicherheit bei Restriktion durch 1.0 (links) und Niveaüfläche des Ersatzmodells (rechts) mit schwarz eingezeichnetem Startpunkt.

Verfahren in blau respektive rot verzeichnet. Wie Abbildung 5.15, rechts, zeigt, handelt es sich, mit Ausnahme des Punktes (0.28, 0.52), um dieselben Minima, die wir ohne Restriktion der Unsicherheit in Abschnitt 5.1.2 erhalten haben.

Abbildung 5.16, links, zeigt die Konvergenzverläufe des Augmented-Lagrange-Verfahrens mit den verschiedenen Verfahren zur unrestringierten Minimierung. Das Fletcher-Reeves- und das Quasi-Newton-Verfahren konvergieren gegen das lokale Minimum 3 bei (0.28, 0.52) mit dem Funktionswert 1.49. Die beiden verbleibenden Verfahren konvergieren hingegen gegen Minimum 1 bei (0.16, 0.49) mit dem Funktionswert 1.64.

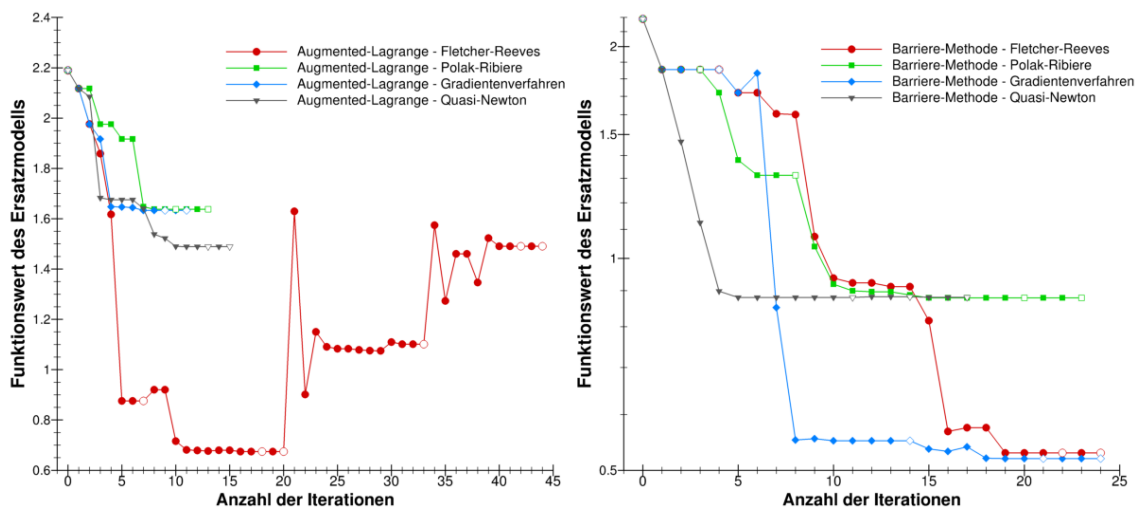


Abbildung 5.16: Konvergenzverläufe des Augmented-Lagrange-Verfahrens (links) und der Barriere-Methode (rechts) bei Restriktion der Unsicherheit durch 1.0 mit verschiedenen Verfahren zur unrestringierten Minimierung (ausgefülltes Symbol: unrestringiertes Verfahren, nichtausgefülltes Symbol: restringiertes Verfahren).

Bei genauerer Betrachtung des Fletcher-Reeves-Konvergenzverlaufs fallen viele Sprünge mit Funktionswerterhöhung auf. Diese treten nicht nur nach einer Augmented-Lagrange-Iteration auf (nichtausgefülltes Symbol im Konvergenzverlauf), sondern auch während der Fletcher-Reeves-Iterationen. Bei einer Augmented-Lagrange-Iteration ist dies beispielsweise aufgrund einer Erhöhung des Strafparameters üblich (vgl. vorherige Abschnitte). Dieses Verhalten während der Fletcher-Reeves-Iterationen mag vorerst verwundern, da das Verfahren nur Punkte als nächste Iterierte annimmt, deren Funktionswert kleiner als jener der vorherigen Iterierten ist. Es lässt sich jedoch durch das Augmented-Lagrange-Verfahren erklären. Das Fletcher-Reeves-Verfahren minimiert nicht das Ersatzmodell, sondern die um Strafterme erweiterte Augmented-Lagrange-Funktion Φ (vgl. Abschnitt 4.2). Betrachten wir die Ausgabewerte des Fletcher-Reeves-Verfahrens genauer, so stellt sich in allen Fällen mit einer Funktionswerterhöhung des Ersatzmodells zugleich eine Reduktion der Nebenbedingungsverletzung $n(x)$ und somit des Funktionswertes $\Phi(x)$ ein.

Die drei verbleibenden Verfahren zeigen das übliche Konvergenzverhalten in einer guten Anzahl an Iterationen. Das Gradientenverfahren mit der geringsten Iterationsanzahl ist hier hervorzuheben. Das Quasi-Newton-Verfahren benötigt zwar länger zur Konvergenz, jedoch gelingt diese, im Gegensatz zum Fletcher-Reeves-Verfahren, in einer akzeptablen Anzahl an Iterationen (vgl. Abb. 5.16, links).

Abbildung 5.16, rechts, zeigt die Konvergenzverläufe der Barriere-Methode bei verschiedenen Verfahren zur unrestringierten Minimierung. Das Fletcher-Reeves- und das Gradientenverfahren konvergieren gegen Minimum 4 bei $x = (0.65, 0.49)$ mit dem Funktionswert 0.52, während das Polak-Ribière- und das Quasi-Newton-Verfahren gegen Minimum 3 bei $x = (0.45, 0.50)$ mit dem Funktionswert 0.88 konvergieren. Das Quasi-Newton-Verfahren weist das beste Konvergenzverhalten mit der geringsten Iterationsanzahl auf. Die drei verbleibenden Verfahren verfügen über eine annähernd gleiche, noch annehmbare Anzahl an Iterationen.

Das Gradientenverfahren zeigt eine Funktionswerterhöhung nach Iteration 6 ähnlich zu denen des Augmented-Lagrange-Verfahrens mit Fletcher-Reeves. Der steigende Funktionswert lässt sich hier durch eine Verringerung der Unsicherheit $h(x)$ begründen.

Vergleichen wir das Augmented-Lagrange-Verfahren und die Barriere-Methode, so zeigt das erstgenannte Verfahren, insbesondere wenn wir das Fletcher-Reeves-Verfahren außen vor lassen, eine schnellere Konvergenz. Das Quasi-Newton-Verfahren zur unrestringierten Minimierung weist wiederum ein stabiles, wenn auch beim Augmented-Lagrange-Verfahren im Vergleich etwas langsames Konvergenzverhalten auf.

Während der letzten beiden Auswertungen (Restriktion durch 0.8 und 1.0) konnten wir, entgegen den vorherigen Untersuchungen, eine bessere Konvergenz des Augmented-Lagrange-Verfahrens im Vergleich zur Barriere-Methode feststellen. Dies ist anscheinend durch die nun auftretende Restriktion der Unsicherheit, mit der die Iterierten konfrontiert sind, begründet. Um diese Aussage zu überprüfen, betrachten wir im Folgenden das Optimierungsproblem mit neuem Startpunkt $(0.852, 0.277)$ und restringieren die Unsicherheit durch den Wert 0.95.

Restriktion der Unsicherheit durch 0.95 bei verändertem Startpunkt

Die Menge der unzulässigen Punkte $\{x \in [0, 1] \times [0, 1] \mid h(x) > 0.95\}$ ist in Abbildung 5.17, links, in dunkelrot gekennzeichnet. Der neue Startpunkt $(0.852, 0.277)$ mit einer ho-

hen Unsicherheit von $h(x) = 0.85$ ist in schwarz, der Endpunkt $(0.65, 0.49)$ aller Verfahren in blau eingezeichnet. Aufgrund der Nähe des Startpunktes zur unzulässigen Menge (vgl. Abb. 5.17, links) können wir mit Hilfe der folgenden Untersuchungen die obige Aussage bezüglich der besseren Konvergenz des Augmented-Lagrange-Verfahrens untermauern.

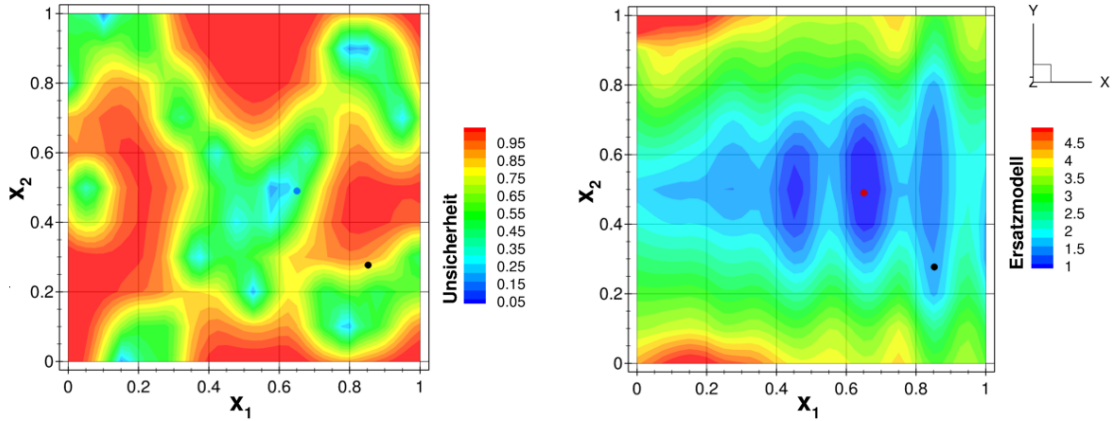


Abbildung 5.17: Niveauefläche der Unsicherheit bei Restriktion durch 0.95 (links) und Niveauefläche des Ersatzmodells (rechts) mit schwarz eingezeichnetem Startpunkt $(0.852, 0.277)$ und blau respektive rot eingezeichnetem Endpunkt $(0.65, 0.49)$ aller Verfahren.

Abbildung 5.18 zeigt die Konvergenzverläufe des Augmented-Lagrange-Verfahrens und der Barriere-Methode im Vergleich. Um Konvergenzunterschiede der unrestringierten Verfahren beim Augmented-Lagrange-Verfahren erkennen zu können, ist eine große Auflösung

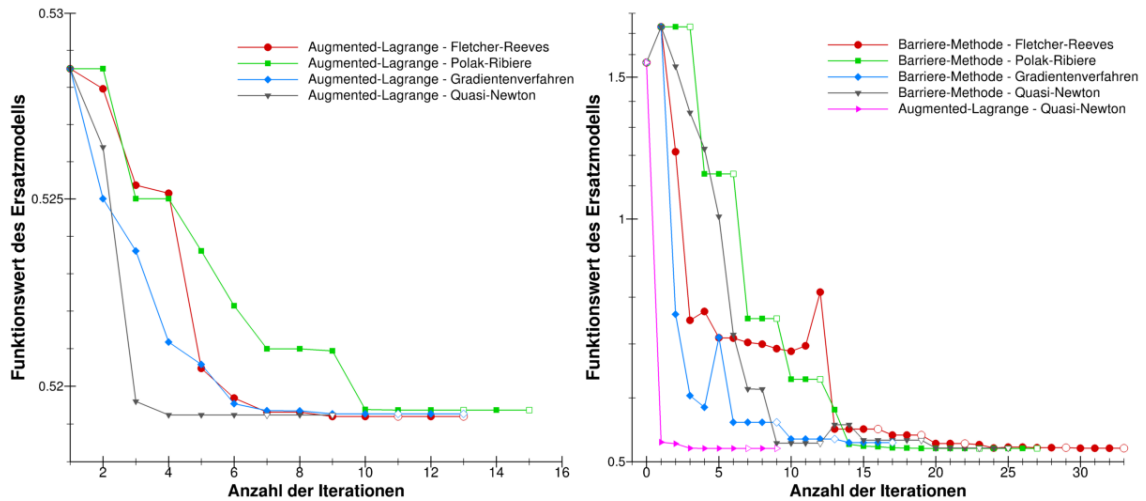


Abbildung 5.18: Konvergenzverläufe des Augmented-Lagrange-Verfahrens (links) und der Barriere-Methode (rechts, zum Vergleich mit Augmented-Lagrange-/Quasi-Newton-Verfahren) bei Restriktion der Unsicherheit durch 0.95 und verändertem Startpunkt $(0.852, 0.277)$ mit verschiedenen Verfahren zur unrestringierten Minimierung (ausgefülltes Symbol: unrestringiertes Verfahren, nichtausgefülltes Symbol: restringiertes Verfahren).

der Ordinatenachse und das Auslassen der ersten Iteration notwendig. Bereits nach dem ersten Schritt in Richtung des Antigradienten befinden sich die Verfahren nahe dem Minimum (siehe Abb. 5.18, links). Das Quasi-Newton-Verfahren zeigt daraufhin das beste Konvergenzverhalten mit der geringsten Iterationsanzahl, allerdings weisen auch die drei verbleibenden Verfahren ein stabiles Konvergenzverhalten mit einer annehmbaren Anzahl an Iterationen auf.

Die Konvergenzverläufe der Barriere-Methode mit verschiedenen Verfahren zur unrestringierten Minimierung sind in Abbildung 5.18, rechts, dargestellt. Fast alle Verfahren weisen Sprünge mit Funktionswerterhöhung auf, welche jedoch stets mit einer Verringerung der Unsicherheit einhergehen. Vergleichen wir die unrestringierten Verfahren, so zeigt das Gradientenverfahren das beste Konvergenzverhalten mit der geringsten Iterationsanzahl. Das Quasi-Newton-Verfahren konvergiert langsamer, allerdings in einer noch akzeptablen Anzahl an Iterationen. Die beiden verbleibenden Verfahren von Fletcher-Reeves und Polak-Ribière benötigen hingegen noch zusätzliche Iterationen zur Konvergenz.

Zum besseren Vergleich der beiden Verfahren zur restringierten Minimierung ist der Konvergenzverlauf des Augmented-Lagrange-/Quasi-Newton-Verfahrens in Abbildung 5.18, rechts, eingezeichnet. Die im Vergleich zu den Verfahren der Barriere-Methode schnellere Konvergenz ist sofort erkennbar. Somit können wir die Aussage bekräftigen, dass das Augmented-Lagrange-Verfahren bei Auftreten *echter* Restriktionen ein besseres Konvergenzverhalten aufweist. Als echte Restriktion verstehen wir in diesem Fall die Nebenbedingung an die Unsicherheit $h(x)$ und nicht wie zuvor nur an die freie Variable x .

Im Vergleich der unrestringierten Verfahren zeigt das Quasi-Newton-Verfahren beim Augmented-Lagrange-Verfahren das beste und bei der Barriere-Methode ein annehmbares Konvergenzverhalten.

5.1.4 Ergebniszusammenfassung der Auswertung der Testfunktion

Im folgenden Abschnitt werden die Ergebnisse der Auswertung der Testfunktion zusammengefasst. In diesem Zusammenhang sei darauf hingewiesen, dass sich die folgenden Aussagen nur auf die Auswertung der Testfunktion stützen und somit nicht notwendigerweise Allgemeingültigkeit besitzen.

Die bisherigen Auswertungen zeigen, dass die im Rahmen dieser Arbeit implementierten Algorithmen nichtlineare Optimierungsprobleme der Form (P) mit nichtlinearen Nebenbedingungen problemlos lösen können. Das Augmented-Lagrange-Verfahren wie auch die Barriere-Methode konvergieren mit allen Verfahren zur unrestringierten Minimierung unabhängig vom Startpunkt gegen ein lokales Minimum.

Ferner können wir festhalten, dass die Barriere-Methode im Fall der Restriktion der freien Variable x ein besseres Konvergenzverhalten zeigt. Wird die Unsicherheit $h(x)$ jedoch zusätzlich restringiert, so weist das Augmented-Lagrange-Verfahren meist eine schnellere Konvergenz auf. Bei Untersuchung der Barriere-Methode mit Restriktion der Unsicherheit lässt sich zudem erkennen, dass aufgrund des logarithmischen Barriere-Terms unüberwindbare Schwellen infolge erhöhter Unsicherheiten entstehen, die jedoch nicht die Restriktion verletzen (vgl. Restriktion durch 0.7 und teilweise 0.8). Beim Augmented-Lagrange-Verfahren tritt dieses nachteilige Verhalten nur in in sehr abgeschwächter Form auf.

Vergleichen wir die unrestringierten Verfahren, so zeigt das Quasi-Newton-Verfahren das stabilste Konvergenzverhalten und mehrheitlich die geringste Anzahl an Iterationen. Ist

dies nicht der Fall, so befindet sich die Iterationsanzahl in einer akzeptablen Größenordnung. Erwähnenswert sei die beim Augmented-Lagrange-Verfahren vereinzelt auftretende zu kleine Startschrittweite, sodass hier eine Anpassung von SIGMA_SHORT notwendig ist.

Entgegen der Vermutung zeigt das Gradientenverfahren teilweise ein sehr gutes Konvergenzverhalten. Andererseits treten Problemstellungen mit überaus langsamer Konvergenz auf, sodass das Gradientenverfahren ungeachtet dessen keine Alternative zum Quasi-Newton-Verfahren darstellt.

Die Verfahren von Fletcher-Reeves und Polak-Ribière zeigen nie ein herausragendes Konvergenzverhalten. Oftmals konvergiert eines der beiden Verfahren langsamer als das Gradientenverfahren. Hier erwartet man aufgrund der Theorie eine schnellere Konvergenz (vgl. Kapitel 4). Diese konnten wir bisher nicht dokumentieren. Das in Harzheim [38] postulierte bessere numerische Verhalten des Polak-Ribière-Verfahrens konnte im Rahmen der vorausgehenden Auswertungen nicht nachgewiesen werden.

Abschnitt 5.1.3 untersucht die Auswirkungen unterschiedlich starker Restriktionen an die Unsicherheit. Wir erkennen, dass insbesondere die Barriere-Methode, aber auch das Augmented-Lagrange-Verfahren, bei zu starker Restriktion der Unsicherheit Gefahr laufen, sich in einer Senke mit niedriger Unsicherheit aber relativ hohem Funktionswert festzulaufer (vgl. Restriktion durch 0.7 und 0.8). Eine gute Vorgehensweise bei der Optimierung könnte eine zu Beginn schwache oder nicht vorhandene Unsicherheitsrestriktion sein, welche je nach Ergebnis entsprechend angepasst respektive verstärkt wird. Weist das Ergebnis schon eine sehr geringe Unsicherheit auf, so kann hier der wahre Funktionswert berechnet und mit dem approximierten Wert des Ersatzmodells verglichen werden. Diese beiden Werte sollten keine außerordentliche Differenz aufweisen. Ist die Unsicherheit jedoch zu groß, so besteht eine erhöhte Wahrscheinlichkeit, dass der wahre Funktionswert von dem durch das Ersatzmodell approximierten Wert abweicht. Diese Abweichung kann zugunsten des Benutzers, wie es bei Minimum 2 oder Minimum 5 der Fall ist (vgl. Abb. 5.2, Mitte), oder zu dessen Ungunsten sein. Welcher der beiden genannten Fälle tatsächlich eintritt, lässt sich nur durch Berechnung des wahren Funktionswertes ermitteln. Ist dieser erheblich größer als der approximierte Wert und für die darauffolgenden Untersuchungen aus diesem Grund unbrauchbar, so ist eine stärkere Restriktion der Unsicherheit notwendig.

5.2 Diskussion anhand eines Beispiels aus dem aerodynamischen Verdichterdesign

Im folgenden Abschnitt untersuchen wir ein reales Optimierungsproblem aus dem aerodynamischen, zweidimensionalen Verdichterprofildesign. Es handelt sich zunächst um eine Aufgabenstellung aus der Mehrzieloptimierung. Die zwei Zielfunktionen beziehungsweise aerodynamischen Größen Totaldruckverhältnis (engl. *total pressure ratio*, kurz TPR)

$$p_{\text{TPR}} = \frac{p_{\text{T, out}}}{p_{\text{T, in}}}$$

und Totaldruckverlust (engl. *total pressure loss*, kurz TPL)

$$p_{\text{TPL}} = \frac{p_{\text{T, out}} - p_{\text{T, in}}}{p_{\text{T, in}} - p_{\text{in}}}$$

sind zu maximieren respektive zu minimieren. Hierbei bezeichnet $p_{\text{T, in}}$ den Totaldruck vor Eintritt in den Verdichter, $p_{\text{T, out}}$ den Totaldruck nach Austritt aus dem Verdichter und p_{in}

den statischen Druck vor Eintritt in den Verdichter. Der Totaldruck setzt sich aus dem statischen und dem dynamischen Druck zusammen. Im Folgenden betrachten wir neben dem Totaldruckverlust das negative Totaldruckverhältnis und minimieren beide Größen. Dieses Mehrzieloptimierungsproblem reduzieren wir in den nachstehenden Abschnitten durch unterschiedliche Herangehensweisen auf ein Problem der Form (P) (vgl. S. 35), welches wir jeweils im Anschluss lösen.

Die zu optimierenden Verdichterprofile werden mit Hilfe der fünf freien Variablen

- Staffelungswinkel β_{ST} ,
- Vorderkantenwinkel β_{LE} (engl. *leading edge*),
- Hinterkantenwinkel β_{TE} (engl. *trailing edge*),
- y -Wert y_{SS} des B-Spline-Kontrollpunktes auf der Saugseite und
- Zuström-Machzahl MA_{in}

beschrieben. Bei den ersten vier Variablen handelt es sich um geometrische Größen zur Beschreibung des Verdichterprofils (vgl. Abb. 5.19), während die letzte Variable eine aerodynamische Größe ist, welche die Geschwindigkeit des einströmenden Gases festlegt.

Die Variable y_{SS} hat unter anderem Einfluss auf die Wölbung des Verdichterprofils: Die Form der Oberseite des Profils (Saugseite) wird durch eine Interpolation mittels B-Splines erzeugt. Die dafür notwendigen Kontrollpunkte, sogenannte De-Boor-Punkte, sind in Abbildung 5.19 blau eingezeichnet. In unserem Fall ist somit der y -Wert des mittleren De-Boor-Punktes variabel und das Profil lässt sich durch Vergrößerung bzw. Verkleinerung von y_{SS} mehr respektive weniger wölben.

Aufgrund der geometrischen Einschränkungen beim Verdichterprofildesign werden die freien Variablen in allen nachstehenden Unterabschnitten wie folgt restringiert:

$$\begin{aligned}
100^\circ &\leq \beta_{ST} \leq 120^\circ \\
120^\circ &\leq \beta_{LE} \leq 150^\circ \\
90^\circ &\leq \beta_{TE} \leq 110^\circ \\
0.1 &\leq y_{SS} \leq 0.3 \\
0.7 &\leq MA_{in} \leq 0.78
\end{aligned} \tag{5.1}$$

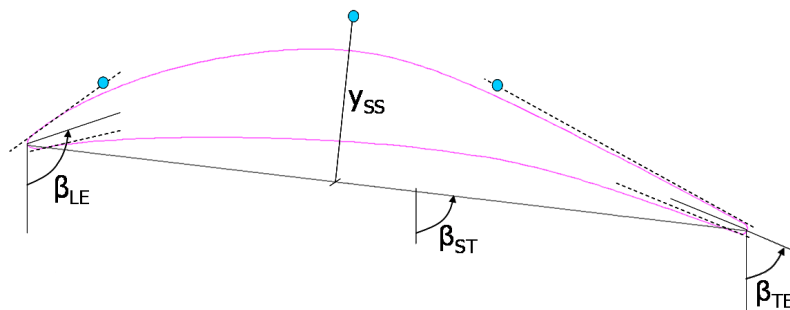


Abbildung 5.19: Skizze eines Verdichterprofils (pink) mit den freien, geometrischen Variablen Staffelungswinkel β_{ST} , Vorderkantenwinkel β_{LE} , Hinterkantenwinkel β_{TE} und y -Wert y_{SS} des B-Spline-Kontrollpunktes.

Da die Machzahl MA_{in} kleiner als eins ist, handelt es sich um eine Unterschall-, eine sogenannte supersonische Strömung (vgl. [5]).

Die Datenbasis zur Erstellung des Ersatzmodells mit Hilfe des Direct Gradient Enhanced Kriging-Verfahrens (Direct GEK, vgl. Abschnitt 3.2.5) besteht aus unterschiedlichen beprobten Punkten beziehungsweise Stützstellen des \mathbb{R}^5 , genauer aus 1206 berechneten Verdichterprofilkonfigurationen. Diese Verdichterprofile stammen aus der Datenbasis (vgl. Abb. 2.3, Mitte unten) einer zuvor am DLR-Institut für Antriebstechnik durchgeführten Optimierung mit AutoOpti auf Grundlage eines genetischen Algorithmus.

Aufgrund der Optimierung von zwei Zielfunktionen gibt es nicht nur eine optimale Verdichterprofilkonfiguration, sondern mehrere, die sogenannte Pareto-Menge. Für eine genaue Definition und eine kurze Einführung in die Mehrzieloptimierung sei auf Anhang A.1 verwiesen. Abbildung 5.20 zeigt die beiden Zielfunktionswerte für alle berechneten Verdichterprofilkonfigurationen respektive Stützstellen und die blau eingezeichnete Pareto-Menge. Da die Datenbasis aus einer zuvor durchgeführten Optimierung mit AutoOpti stammt, besitzen viele dieser Verdichterprofile schon vor Beginn unserer Optimierung mit Gradienten-basierten Verfahren relativ niedrige Zielfunktionswerte (vgl. Abb. 5.20). Ziel dieses Abschnitts wird sein, sich mit Hilfe der Gradienten-basierten Optimierungsverfahren diesen schon berechneten Pareto-Optima anzunähern und wenn möglich diese zu erreichen.

Die Aufstellung der Datenbasis, also die Berechnung der beiden Zielfunktionswerte für die 1206 Verdichterprofilkonfigurationen, erfolgte mit dem Strömungslöser MISES (vgl. [47]). Dieser löst die Euler-Gleichung zur Beschreibung des Strömungsverhaltens von reibungsfreien Fluiden. MISES ist gekoppelt mit einem Grenzschichtmodell, welches die turbulenten Transportvorgänge in der Grenzschicht simuliert.

Im DLR und im Rahmen dieser Arbeit wird MISES ausschließlich als „Black Box“ benutzt. Wir entschieden uns für die Verwendung von MISES statt TRACE, da beide Strömungslöser die gesuchten Zielfunktionswerte Totaldruckverhältnis und Totaldruckver-

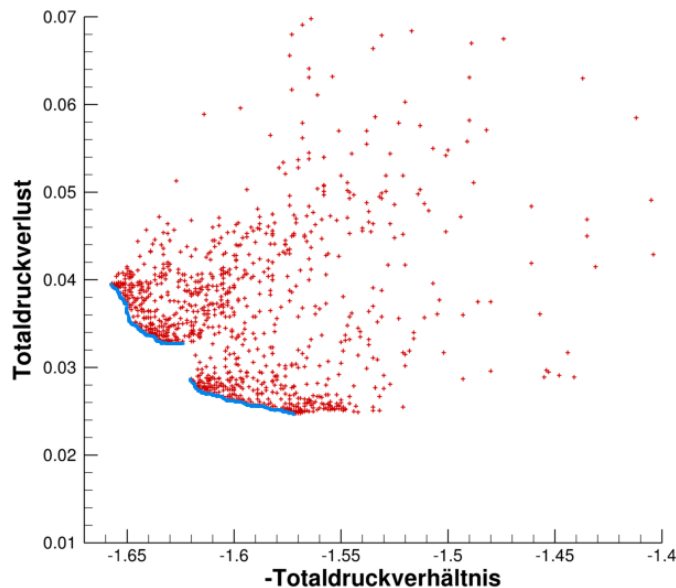


Abbildung 5.20: Zielfunktionswerte aller mit MISES berechneten Verdichterprofilkonfigurationen (Datenbasis des Ersatzmodells) in rot und Pareto-Menge in blau.

lust berechnen können. Aufgrund der Euler-Gleichung, welche ein Spezialfall der Navier-Stokes-Gleichung ist, berechnet MISES diese Größen jedoch in kürzerer Zeit als der Strömungslöser TRACE, welcher die komplexeren RANS-Gleichungen löst. Die Genauigkeit der MISES-Lösungen ist für unsere Testzwecke ausreichend gut.

Nach Abschluss der Gradienten-basierten Optimierung des Ersatzmodells werden vielversprechende Punkte beziehungsweise Verdichterprofilkonfigurationen mit MISES nachgerechnet. Dies ist notwendig, da die Resultate aufgrund der Optimierung des Ersatzmodells, also einer Approximation der wahren Funktionen *Totaldruckverhältnis* und *Totaldruckverlust*, mit einer Unsicherheit versehen sind. Wir interessieren uns jedoch letztlich für das mit Hilfe von MISES berechnete wahre Totaldruckverhältnis und den wahren Totaldruckverlust dieser vielversprechenden Verdichterprofilkonfigurationen.

In den folgenden Abschnitten werden wir vorerst das Mehrzieloptimierungsproblem durch Betrachtung einer gewichteten Summe der Zielfunktionen auf ein herkömmliches Optimierungsproblem mit einer Zielfunktion ohne Restriktion der Unsicherheiten reduzieren. Im Anschluss untersuchen wir die Auswirkungen einer Maximierung des Totaldruckverhältnisses unter Restriktion des Totaldruckverlustes. Schließlich betrachten wir erneut eine gewichtete Summe der Zielfunktionen und analysieren die Ergebnisse bei unterschiedlich starken Restriktionen der Unsicherheiten.

Im Folgenden seien die Parameter des Armijo-Verfahrens wie in Abschnitt 4.3.5 angegeben gewählt. Für die restringierten Verfahren findet die prozentuale Abbruchbedingung mit $p_1 = 1$, $p_2 = 0.05$, $p_3 = 1$ und $p_4 = 0.05$ Anwendung (vgl. S. 37). Die Abbruchtoleranzen ω und ω_x der unrestringierten Verfahren setzen wir auf die Werte 10^{-8} respektive $12 \cdot 10^{-9}$ (vgl. S. 40). Diese Parameter gelten, sofern keine anderen Werte angegeben sind, für alle folgenden Unterabschnitte.

5.2.1 Optimierung der gewichteten Summe ohne Restriktion

Im folgenden Abschnitt reduzieren wir die obige Mehrzieloptimierungsaufgabe durch Betrachtung einer gewichteten Summe der beiden Zielfunktionen auf ein Problem der Form (P) ohne nichtlineare Nebenbedingungen (vgl. S. 35). Die freien Variablen werden wie in (5.1) gefordert restringiert. Im Laufe des Abschnitts werden verschiedene Gewichtungen der beiden Zielfunktionen betrachtet.

Optimierung von $-0.628 \cdot p_{\text{TPR}} + 27.027 \cdot p_{\text{TPL}}$ ohne Restriktion

Zunächst gewichten wir die beiden mit Hilfe der Ersatzmodelle approximierten Zielfunktionen *Totaldruckverhältnis* p_{TPR} und *Totaldruckverlust* p_{TPL} in natürlicher Weise durch das arithmetische Mittel der Funktionswerte aller 1206 in der Datenbasis enthaltenen Verdichterprofile. Somit erhalten wir für das Totaldruckverhältnis den Faktor $-\frac{1}{1.593} \approx -0.628$ sowie den Faktor $\frac{1}{0.037} \approx 27.027$ für den Totaldruckverlust. Die Unsicherheit wird, wie in Abschnitt 5.1.4 vorgeschlagen, zunächst nicht restringiert.

Abbildung 5.21, links, zeigt die mit Hilfe des Direct GEK berechneten approximierten Zielfunktionswerte für alle konvergierten Verfahren. Das Gradientenverfahren überschreitet bei allen nachfolgenden Untersuchungen die maximale Iterationsanzahl von 250 Iterationen für unrestringierte Verfahren (vgl. Abb. 5.22). Somit entfällt im Folgenden sowohl für das Augmented-Lagrange-Verfahren als auch für die Barriere-Methode die Analyse für dieses Verfahren zur unrestringierten Minimierung.

Die beiden Zielfunktionswerte des Startpunktes

$$x^{(0)} = (114.9, 136, 100, 0.2170, 0.7447)^T$$

sind in Abbildung 5.21 blau gekennzeichnet. Die Barriere-Methode konvergiert mit allen Verfahren zur unrestringierten Minimierung gegen $\tilde{x}^{(1)} = (113.13, 137.22, 99.99, 0.26, 0.72)^T$ mit dem approximierten negativen Totaldruckverhältnis -1.48 und dem Totaldruckverlust 0.0137 (vgl. Abb. 5.21, links). Selbiges Resultat erhalten wir beim Augmented-Lagrange-/Fletcher-Reeves-Verfahren. Die Einträge von $\tilde{x}^{(1)}$ berühren nicht die Grenzen (5.1) der freien Variablen. Des Weiteren ist die Norm des Gradienten der gewichteten Zielfunktion $-0.628 \cdot p_{\text{TPR}} + 27.027 \cdot p_{\text{TPL}}$ bei $\tilde{x}^{(1)}$ klein. Somit können wir davon ausgehen, dass die Verfahren tatsächlich gegen ein lokales Minimum der gewichteten Zielfunktion konvergieren.

Das Augmented-Lagrange-Verfahren mit dem Quasi-Newton-Verfahren zur unrestringierten Minimierung konvergiert gegen den Punkt $\tilde{x}^{(2)} = (113.12, 137.22, 90.00, 0.26, 0.72)^T$ mit einem im Vergleich zu $\tilde{x}^{(1)}$ verbesserten approximierten Totaldruckverhältnis von 1.49 (vgl. Abb. 5.21, links). Der Wert des approximierten Totaldruckverlustes ist unverändert 0.0137. Der Punkt $\tilde{x}^{(2)}$ unterscheidet sich nur im dritten Eintrag, also beim Hinterkantenwinkel β_{TE} , signifikant von $\tilde{x}^{(1)}$. Auch in diesem Fall besitzt der Gradient der gewichteten Zielfunktion bei $\tilde{x}^{(2)}$ eine kleine Norm; damit handelt es sich bei diesem Punkt ebenfalls um ein lokales Minimum. Das Augmented-Lagrange-Verfahren mit Polak-Ribière zur unrestringierten Minimierung konvergiert gegen ein weiteres lokales Minimum bei dem Punkt $\tilde{x}^{(3)} = (114.76, 136.18, 99.99, 0.26, 0.72)^T$. Beide Funktionswerte verbessern sich zwar gegenüber denen des Startpunktes, allerdings liegt $\tilde{x}^{(3)}$ mit einem approximierten Totaldruckverhältnis von -1.47 und einem approximierten Totaldruckverlust von 0.037 deutlich hinter den beiden vorhergenannten Punkten (vgl. Abb. 5.21, links). Aus Sicht der implementierten Verfahren besteht hier jedoch kein Unterschied, da diese gegen lokale und nicht globale Minima konvergieren.

Abbildung 5.22 zeigt die Konvergenzverläufe des Augmented-Lagrange-Verfahrens und der Barriere-Methode. Wir registrieren eine im Vergleich zu den vorher untersuchten Optimie-

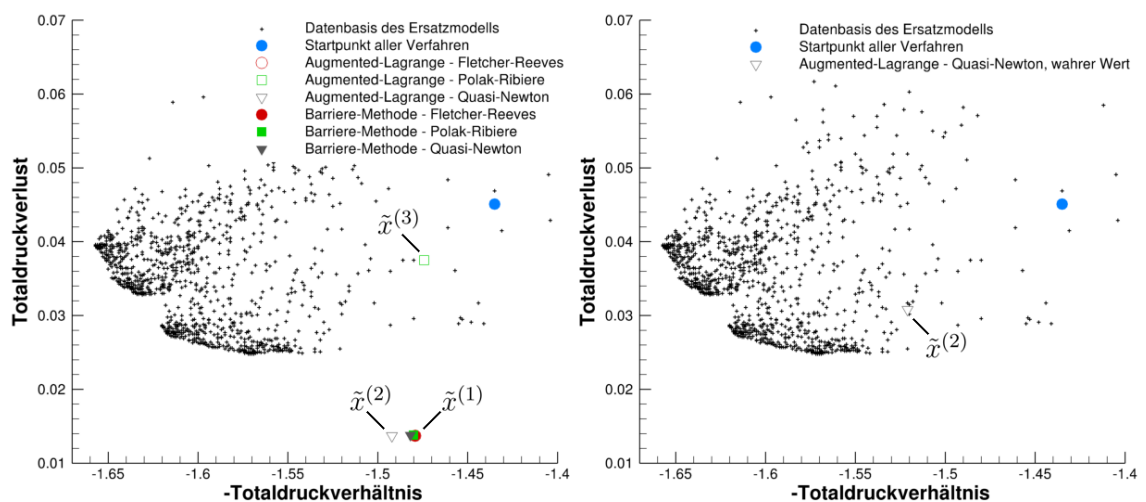


Abbildung 5.21: Zielfunktionswerte der Ersatzmodell-Datenbasis in schwarz, Approximation der Zielfunktionswerte aller Endpunkte der konvergierten Verfahren (links) sowie wahre Werte des Augmented-Lagrange-/Quasi-Newton-Endpunktes (rechts) bei Gewichtung mit den Faktoren -0.628 und 27.027.

rungsproblemen in Abschnitt 5.1 erhöhte Anzahl an Iterationen bei allen Verfahren. Diese liegt in der nun höheren Dimension des Definitionsbereichs begründet.

Vergleichen wir die unrestringierten Verfahren, so weist das Quasi-Newton-Verfahren in beiden Fällen die schnellste Konvergenz mit der geringsten Anzahl an Iterationen auf. Das Fletcher-Reeves- und das Polak-Ribière-Verfahren unterscheiden sich hier stark. Letzteres zeigt sowohl beim Augmented-Lagrange-Verfahren als auch bei der Barriere-Methode ein besseres numerisches Verhalten und eine geringere Anzahl an Iterationen. Das Gradientenverfahren konvergiert äußerst langsam und bricht nach 250 Iterationen aufgrund des Erreichens der maximalen Iterationsanzahl ab.

Bei den restringierten Verfahren zeigt sich kein eindeutiges Bild. Beim Augmented-Lagrange-Verfahren konvergiert das Quasi-Newton-Verfahren am schnellsten, allerdings zeigen die Verfahren von Fletcher-Reeves und Polak-Ribière bei der Barriere-Methode ein besseres Konvergenzverhalten.

Im Anschluss an die Optimierung des Ersatzmodells berechnen wir mit Hilfe des Strömungslösers MISES die wahren Funktionswerte für den besten Endpunkt $\tilde{x}^{(2)}$ des Augmented-Lagrange-/Quasi-Newton-Verfahrens. Dieser ist, basierend auf den approximierten Werten, aufgrund des sehr geringen approximierten Totaldruckverlustes ein Pareto-Optimum und folglich für unsere Auswertung interessant.

Wie Abbildung 5.21, rechts, zeigt, befinden sich die wahren Zielfunktionswerte von $\tilde{x}^{(2)}$ mit einem Totaldruckverhältnis von 1.521 und einem Totaldruckverlust von 0.0308 weit von den durch das Ersatzmodell vorhergesagten Funktionswerten entfernt. Das Totaldruckverhältnis hat sich verbessert, während sich der Totaldruckverlust so weit verschlechtert hat, dass es sich bei $\tilde{x}^{(2)}$ nicht länger um ein Pareto-Optimum handelt (vgl. Abb. 5.21, rechts). Die Unsicherheiten $h_1(\tilde{x}^{(2)}) = 0.15$ und $h_2(\tilde{x}^{(2)}) = 0.13$ des Totaldruckverhältnisses respektive Totaldruckverlustes sind zu groß, um eine gute Vorhersage der wahren Funktionswerte durch das Ersatzmodell zu gewährleisten. Hier ist eine Restriktion der Unsicherheit notwendig (vgl. Abschnitt 5.2.3).

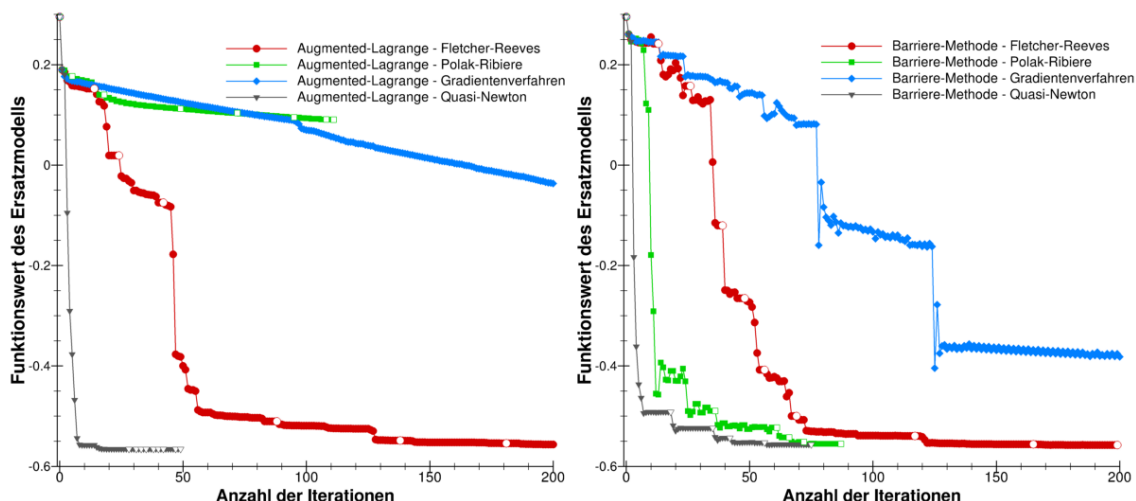


Abbildung 5.22: Konvergenzverhalten des Augmented-Lagrange-Verfahrens (links) und der Barriere-Methode (rechts) bei Gewichtung mit den Faktoren -0.628 und 27.027 und verschiedenen Verfahren zur unrestringierten Minimierung (ausgefülltes Symbol: unrestringiertes Verfahren, nichtausgefülltes Symbol: restringiertes Verfahren).

Zunächst soll jedoch die Gewichtung so angepasst werden, dass das bisher relativ niedrige zu maximierende Totaldruckverhältnis priorisiert wird.

Optimierung von $-1.046 \cdot p_{\text{TPR}} + 9.009 \cdot p_{\text{TPL}}$ ohne Restriktion

In diesem Abschnitt sollen die Auswirkungen einer stärkeren Gewichtung des Totaldruckverhältnisses untersucht werden. Im Folgenden wählen wir für dieses den Gewichtungsfaktor $-\frac{1}{1.593} \cdot \frac{5}{3} \approx -1.046$ sowie für den Totaldruckverlust den Faktor $\frac{1}{0.037} \cdot \frac{1}{3} \approx 9.009$. Die Unsicherheit wird nicht restringiert und der Startpunkt ist weiterhin der Punkt $x^{(0)} = (114.9, 136, 100, 0.2170, 0.7447)^T$.

Abbildung 5.23, links, zeigt die approximierten Zielfunktionswerte der Endpunkte aller konvergierten Verfahren. Wir erkennen, dass die stärkere Gewichtung des Totaldruckverhältnisses eine teilweise Verschiebung der approximierten Funktionswerte nach links zur Folge hat (siehe Funktionswerte der Barriere-Methode mit Polak-Ribière- oder auch Quasi-Newton-Verfahren). Die restlichen Verfahren konvergieren jedoch weiterhin gegen Punkte nahe den im vorherigen Abschnitt untersuchten Punkten $\tilde{x}^{(1)}$ und $\tilde{x}^{(2)}$.

So konvergiert etwa das Augmented-Lagrange-/Polak-Ribière-Verfahren gegen das lokale Minimum bei $\tilde{x}^{(1)}$. Einen nahe $\tilde{x}^{(1)}$ liegenden Endpunkt besitzen sowohl das Augmented-Lagrange-Verfahren als auch die Barriere-Methode mit dem Fletcher-Reeves-Verfahren zur unrestringierten Minimierung. Das Augmented-Lagrange-/Quasi-Newton-Verfahren konvergiert erneut gegen $\tilde{x}^{(2)}$.

Bei den beiden neu erreichten Endpunkten handelt es sich um den Punkt $\tilde{x}^{(4)} = (110.92, 137.90, 99.81, 0.25, 0.74)^T$ mit dem approximierten Totaldruckverhältnis 1.52 und dem approximierten Totaldruckverlust 0.0204 sowie um den Punkt $\tilde{x}^{(5)} = (108.18, 137.48, 99.60, 0.21, 0.74)^T$ mit den angenäherten Funktionswerten -1.52 respektive 0.0293 (vgl. Abb. 5.23, links). Die Norm des Gradienten der gewichteten Zielfunktion ist in beiden Punkten gering.

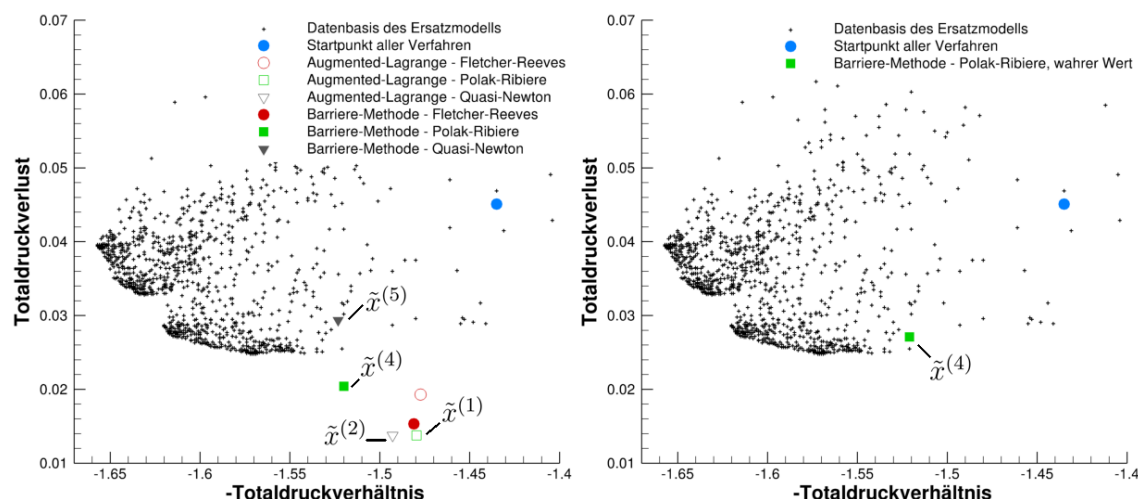


Abbildung 5.23: Zielfunktionswerte der Ersatzmodell-Datenbasis in schwarz, Approximation der Zielfunktionswerte aller Endpunkte der konvergierten Verfahren (links) sowie wahre Werte der Barriere-Methode mit Polak-Ribière-Verfahren (rechts) bei Gewichtung mit den Faktoren -1.046 und 9.009.

Die Konvergenzverläufe des Augmented-Lagrange-Verfahrens und der Barriere-Methode sind in Abbildung 5.24 dargestellt. Das Gradientenverfahren konvergiert wiederum sehr langsam und überschreitet bei beiden Verfahren die maximale Iterationsanzahl von 250 Iterationen.

Sowohl beim Augmented-Lagrange-Verfahren als auch bei der Barriere-Methode weist das Quasi-Newton-Verfahren die geringste Anzahl an Iterationen auf. Das in Abbildung 5.24, rechts, dargestellte anfängliche Verhalten des Quasi-Newton-Verfahrens bei der Barriere-Methode liegt in den Restriktionen der freien Variablen begründet, welche in den logarithmischen Barriere-Term eingehen. Der Wert des logarithmischen Barriere-Terms der zu minimierenden Funktion Φ verringert sich, sobald sich die Iterierte von den Grenzen (5.1) entfernt und weiter in die Mitte des zulässigen Bereichs bewegt. Dies kann jedoch eine Erhöhung des Funktionswertes der gewichteten Zielfunktion, wie dies zwischen Iteration 5 und Iteration 11 der Fall ist, zur Folge haben.

Die Verfahren von Fletcher-Reeves und Polak-Ribière weisen je nach Art des restringierten Verfahrens ein unterschiedliches Verhalten auf. So zeigt die erstgenannte Methode beim Augmented-Lagrange-Verfahren ein besseres Verhalten, während bei der Barriere-Methode das Polak-Ribière-Verfahren in einer geringeren Anzahl an Iterationen konvergiert.

Vergleichen wir die restringierten Verfahren, so weist die Barriere-Methode im Vergleich zum Augmented-Lagrange-Verfahren im Mittel eine geringfügig höhere Anzahl an Iterationen auf. Das beste und stabilste Konvergenzverhalten zeigt das Augmented-Lagrange-/Quasi-Newton-Verfahren.

Nach Optimierung des Ersatzmodells werden sodann mit Hilfe des Strömungslösers MISES die wahren Funktionswerte für das nach den approximierten Werten zu urteilende neu ermittelte Pareto-Optimum $\tilde{x}^{(4)}$ berechnet. Die wahren Funktionswerte für den nach Abbildung 5.23, links, zweiten vielversprechenden Punkt $\tilde{x}^{(2)}$ wurden bereits im vorherigen Abschnitt errechnet und diskutiert (vgl. Abb. 5.21, rechts).

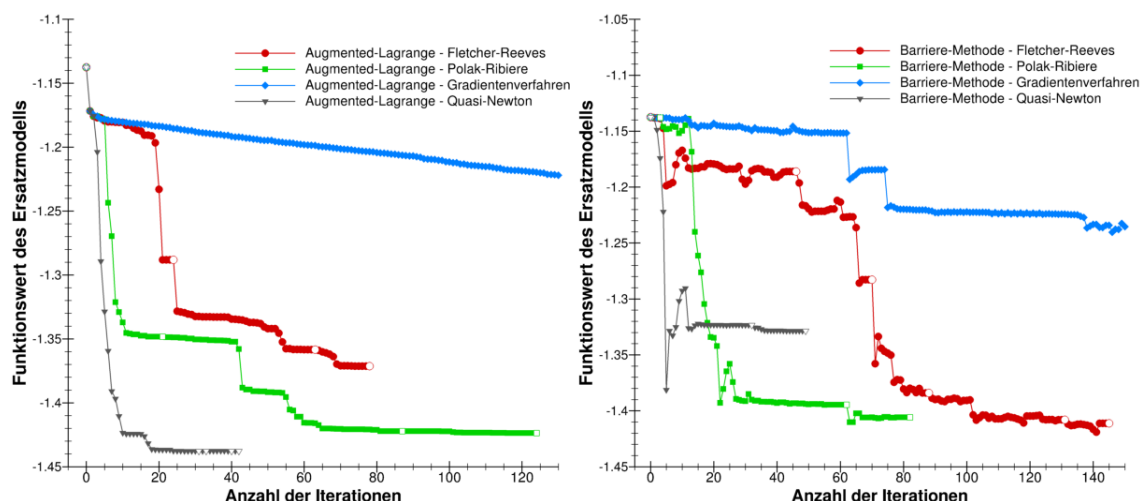


Abbildung 5.24: Konvergenzverhalten des Augmented-Lagrange-Verfahrens (links) und der Barriere-Methode (rechts) bei Gewichtung mit den Faktoren -1.046 und 9.009 und verschiedenen Verfahren zur unrestringierten Minimierung (ausgefülltes Symbol: unrestringiertes Verfahren, nichtausgefülltes Symbol: restringiertes Verfahren).

Abbildung 5.23, rechts, zeigt die wahren Funktionswerte für den Endpunkt $\tilde{x}^{(4)}$ der Barriere-Methode mit Polak-Ribière-Verfahren. MISES berechnet ein Totaldruckverhältnis von 1.521 und einen Totaldruckverlust von 0.027. Beim Vergleich der approximierten und wahren Funktionswerte erkennen wir, dass das vorhergesagte und das wahre Totaldruckverhältnis sehr gut übereinstimmen. Einzig der Totaldruckverlust wurde mit einem Wert von 0.0204 besser vorhergesagt.

Der Grund für die sehr gute Vorhersage des Totaldruckverhältnisses kann jedoch nicht mit Hilfe der Unsicherheiten bei $\tilde{x}^{(4)}$, welche das Direct GEK zusätzlich zu den approximierten Funktionswerten liefert, begründet werden: Die Unsicherheit für das Totaldruckverhältnis ist durch $h_1(\tilde{x}^{(4)}) = 0.15$ gegeben, wohingegen die Unsicherheit des Totaldruckverlustes durch $h_2(\tilde{x}^{(4)}) = 0.13$ gegeben ist. Somit besitzt das gut vorhergesagte Totaldruckverhältnis eine größere Unsicherheit als der Totaldruckverlust. Eventuell ist die sehr gute Vorhersage des Totaldruckverhältnisses ein eher zufälliges Resultat des Direct GEK.

Aufgrund des schlechteren wahren Totaldruckverlustes handelt es sich bei $\tilde{x}^{(4)}$ daraufhin nicht mehr um ein Pareto-Optimum.

Im folgenden Abschnitt werden wir die Gewichtung des Totaldruckverhältnisses weiter erhöhen, da vier der sechs im obigen Abschnitt analysierten Verfahren gegen die schon vorher bekannten Punkte $\tilde{x}^{(1)}$ und $\tilde{x}^{(2)}$ mit einem geringen Totaldruckverhältnis konvergieren (vgl. Abb. 5.23, links).

Optimierung von $-1.245 \cdot p_{\text{TPR}} + 0.452 \cdot p_{\text{TPL}}$ ohne Restriktion

In diesem Abschnitt werden die Auswirkungen einer noch stärkeren Gewichtung des Totaldruckverhältnisses durch den Faktor -1.245 untersucht. Der Totaldruckverlust geht durch den Gewichtungsfaktor 0.452 in die gewichtete Summe ein. Der Startpunkt ist weiterhin der Punkt $x^{(0)} = (114.9, 136, 100, 0.2170, 0.7447)^T$ und die Unsicherheit wird im Folgenden nicht restringiert.

Die approximierten Zielfunktionswerte der Endpunkte aller konvergierten Verfahren sind in Abbildung 5.25, links, dargestellt. Die große Gewichtung des Totaldruckverhältnisses bewirkt eine im Vergleich zum vorherigen Abschnitt noch stärkere Verschiebung der approximierten Funktionswerte der Endpunkte nach links. Die geringe Gewichtung des Totaldruckverlustes hat jedoch zur Folge, dass die Endpunkte aller Verfahren, mit Ausnahme des Augmented-Lagrange-/Quasi-Newton-Verfahrens, einen sehr hohen approximierten Totaldruckverlust aufweisen. Anscheinend geht dieser bei der Gewichtung nun in zu geringem Maße ein (vgl. Abb. 5.25, links).

Die unterschiedlichen Verfahren konvergieren gegen die in Abbildung 5.25, links, gekennzeichneten lokalen Minima

$$\begin{aligned}\tilde{x}^{(6)} &= (114.17, 136.21, 99.76, 0.29, 0.78)^T \text{ mit } p_{\text{TPR}} = 1.51 \text{ und } p_{\text{TPL}} = 0.057, \\ \tilde{x}^{(7)} &= (107.34, 141.56, 97.97, 0.29, 0.78)^T \text{ mit } p_{\text{TPR}} = 1.59 \text{ und } p_{\text{TPL}} = 0.034, \\ \tilde{x}^{(8)} &= (108.89, 136.27, 90.00, 0.30, 0.78)^T \text{ mit } p_{\text{TPR}} = 1.62 \text{ und } p_{\text{TPL}} = 0.012, \\ \tilde{x}^{(9)} &= (114.04, 136.26, 99.94, 0.21, 0.74)^T \text{ mit } p_{\text{TPR}} = 1.48 \text{ und } p_{\text{TPL}} = 0.018 \text{ sowie} \\ \tilde{x}^{(10)} &= (109.04, 137.51, 99.56, 0.21, 0.74)^T \text{ mit } p_{\text{TPR}} = 1.52 \text{ und } p_{\text{TPL}} = 0.032.\end{aligned}$$

Die Norm des Gradienten der gewichteten Zielfunktion ist in allen Punkten mit Ausnahme von $\tilde{x}^{(8)}$ klein. Bei diesem zum Augmented-Lagrange-/Quasi-Newton-Verfahren

gehörenden Endpunkt besitzt der Gradient der gewichteten Zielfunktion im vierten Eintrag der zugehörigen freien Variable y_{SS} einen vergleichsweise großen Wert. Dieser erklärt sich durch das Erreichen der oberen Grenze von y_{SS} . Aus demselben Grund tritt das in Abbildung 5.26, links, dargestellte Konvergenzverhalten des Augmented-Lagrange-/Quasi-Newton-Verfahrens nach den Iterationen 86, 103, 120 und 138 auf. Hier ist das typische Verhalten aufgrund einer Unzulässigkeit der Iterierten erkennbar. Erst eine Erhöhung des Strafparameters μ_k führt zur Zulässigkeit der Iterierten im vierten Eintrag. Dies hat jedoch eine Vergrößerung des Zielfunktionswertes zur Folge.

Vergleichen wir die unrestringierten Verfahren, so zeigt das Quasi-Newton-Verfahren bei der Barriere-Methode das schnellste Konvergenzverhalten. Beim Augmented-Lagrange-Verfahren weisen sowohl das Quasi-Newton- als auch das Fletcher-Reeves-Verfahren die geringste Anzahl an Iterationen auf.

Die Verfahren von Fletcher-Reeves und Polak-Ribière zeigen ein unterschiedliches Verhalten je nach Art des restringierten Verfahrens. Ersteres Verfahren zeigt eine geringere Iterationsanzahl beim Augmented-Lagrange-Verfahren, während letzteres eine schnellere Konvergenz bei der Barriere-Methode aufweist. In diesem Zusammenhang sei darauf hingewiesen, dass das Polak-Ribière-Verfahren zwar in beiden Fällen gegen einen Punkt mit geringerem Zielfunktionswert konvergiert, dies jedoch beim Vergleich beider Verfahren nicht maßgeblich ist, da diese gegen lokale und nicht globale Minima konvergieren.

Das Gradientenverfahren wird bei dieser und allen folgenden Auswertungen nicht weiter betrachtet, da es in sämtlichen untersuchten Fällen die maximale Iterationsanzahl der unrestringierten Verfahren von 250 Iterationen überschreitet und bis zum Abbruch ein äußerst langsames Konvergenzverhalten aufweist.

Beim Vergleich der restringierten Verfahren zeigt die Barriere-Methode bei aktueller Gewichtung eindeutig das bessere Konvergenzverhalten (vgl. Abb. 5.26). Diese konvergiert mit allen Verfahren zur unrestringierten Optimierung in deutlich weniger Iterationen.

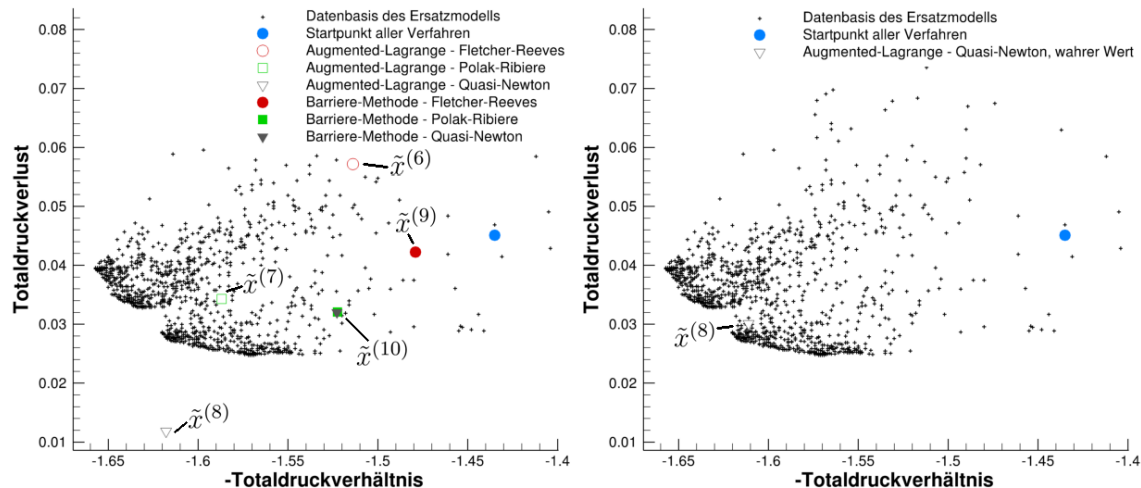


Abbildung 5.25: Zielfunktionswerte der Ersatzmodell-Datenbasis in schwarz, Approximation der Zielfunktionswerte aller Endpunkte der konvergierten Verfahren (links) sowie wahre Werte des Augmented-Lagrange-/Quasi-Newton-Verfahrens (rechts) bei Gewichtung mit den Faktoren -1.245 und 0.452.

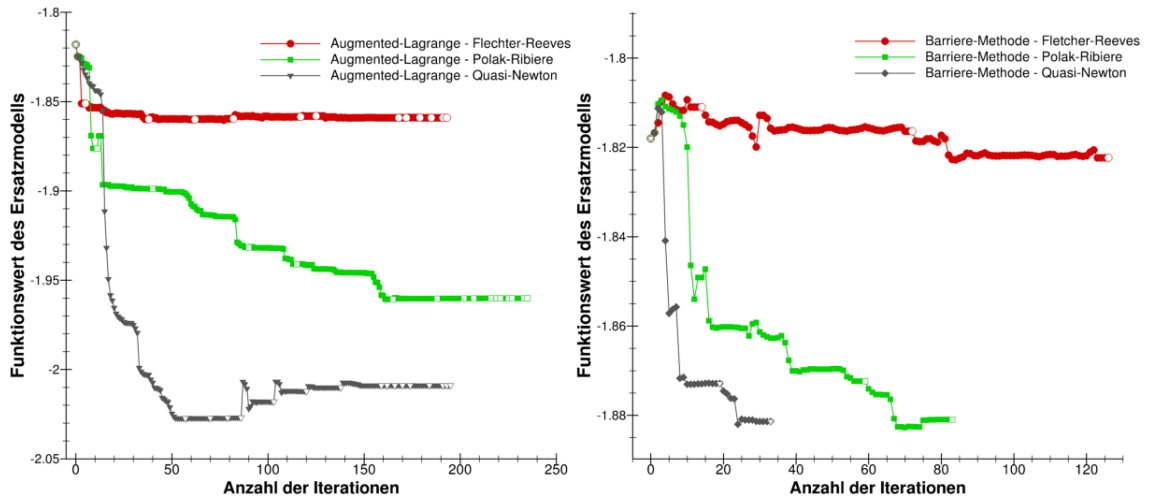


Abbildung 5.26: Konvergenzverhalten des Augmented-Lagrange-Verfahrens (links) und der Barriere-Methode (rechts) bei Gewichtung mit den Faktoren -1.245 und 0.452 und verschiedenen Verfahren zur unrestringierten Minimierung (ausgefülltes Symbol: unrestringiertes Verfahren, nichtausgefülltes Symbol: restringiertes Verfahren).

Im Anschluss an die Optimierung berechnen wir wiederum mit Hilfe des Strömungslösers MISES die wahren Funktionswerte für den vielversprechenden Endpunkt $\tilde{x}^{(8)}$ des Augmented-Lagrange-/Quasi-Newton-Verfahrens (siehe Abb. 5.25, links). Abbildung 5.25, rechts, zeigt diese wahren Funktionswerte. MISES berechnet ein Totaldruckverhältnis von 1.611 und einen Totaldruckverlust von 0.0301 gegenüber den vorhergesagten Werten 1.618 respektive 0.0118. Das Totaldruckverhältnis mit einer Unsicherheit $h_1(\tilde{x}^{(8)}) = 0.069$ wird somit wiederum sehr gut vorhergesagt, der Totaldruckverlust mit $h_2(\tilde{x}^{(8)}) = 0.068$ hingegen falsch. Wie schon im vorherigen Abschnitt lässt die Größe der Unsicherheit keinen Rückschluss auf die Richtigkeit der Vorhersage durch das Direct GEK zu.

Der gesamte vorherige Abschnitt zeigt, dass die Gradienten-basierte Optimierung des Ersatzmodells eines Verdichterdesignproblems mit Hilfe der im Rahmen dieser Arbeit implementierten Verfahren funktioniert. Mit Ausnahme des Gradientenverfahrens, dessen langsames Konvergenzverhalten in der Literatur wohlbekannt ist (vgl. [35], S. 99), konvergieren alle Verfahren gegen lokale Minima der jeweiligen gewichteten Zielfunktion. Die Punkte, gegen die die Verfahren konvergieren, besitzen teilweise sehr vielversprechende approximierte Funktionswerte.

Bei der anschließenden Berechnung mit MISES wichen die wahren Werte von den approximierte Werten oft erkennbar ab. Eine Restriktion der Unsicherheit kann hier eventuell bessere Resultate liefern. Leider konnten wir im vorherigen Abschnitt aus den dort auftretenden Unsicherheiten keinen zuverlässigen Schluss auf die notwendige Größe dieser Restriktion ziehen: Richtig vorhergesagte Werte besaßen teilweise eine höhere Unsicherheit als die falsch approximierte Funktionswerte und Funktionswerte mit einer Unsicherheit von 0.13 wurden genauso schlecht vorhergesagt wie Funktionswerte mit einer um die Hälfte verringerten Unsicherheit von 0.07. Eventuell ist die Restriktion der Unsicherheit die falsche Herangehensweise, da das Direct GEK hier anscheinend zu ungenaue Werte liefert.

Weiter erkennen wir, dass die Gewichtung der beiden Zielfunktionen und die Anpassung dieser Gewichtungsfaktoren zwar die richtigen Tendenzen zeigen, es jedoch sehr schwer kontrollierbar ist, gegen welchen Punkt die einzelnen Verfahren letztendlich konvergieren. Der folgende Abschnitt wird eine Herangehensweise aufzeigen, die bessere Ergebnisse liefert. Vorerst werden wir nicht die Unsicherheit restringieren, sondern die Minimierung des Totaldruckverhältnisses bei unterschiedlich starker Restriktion des Totaldruckverlustes untersuchen.

5.2.2 Optimierung des Totaldruckverhältnisses unter Restriktion des Totaldruckverlustes

Im folgenden Abschnitt betrachten wir die Auswirkungen einer Maximierung des Totaldruckverhältnisses bei Restriktion des Totaldruckverlustes. Es werden unterschiedlich starke Restriktionen untersucht. Insbesondere wollen wir analysieren, ob es möglich ist, die Lücke der Pareto-Front bei einem Totaldruckverhältnis von 1.62 (vgl. Abb. 5.20) mit Hilfe geeigneter Punkte beziehungsweise Verdichterprofilkonfigurationen zu schließen.

Restriktion des Totaldruckverlustes durch 0.04

Im Folgenden soll der Totaldruckverlust durch 0.04 nach oben beschränkt werden, d.h. $p_{\text{TPL}} \leq 0.04$. Bei der Barriere-Methode muss der Startpunkt in der zulässigen Menge liegen (vgl. Verfahren 4.1). Da der bisherige Startpunkt mit einem Totaldruckverlust von 0.0451 die Nebenbedingung an diesen nicht erfüllt (siehe bspw. Abb. 5.25), verwenden wir $x^{(0)} = (110.204, 146.327, 95.306, 0.239, 0.720)^T$ als neuen Startpunkt (siehe Abb. 5.27). Obwohl das Augmented-Lagrange-Verfahren keinen zulässigen Startpunkt benötigt, verwenden wir aus Gründen der besseren Vergleichbarkeit der Resultate den neuen Startpunkt für beide Verfahren der restringierten Minimierung.

Abbildung 5.27, links, zeigt die erreichten Endpunkte des Augmented-Lagrange-Verfahrens und der Barriere-Methode. Aufgrund der Nähe zur Pareto-Front ist insbesondere der End-

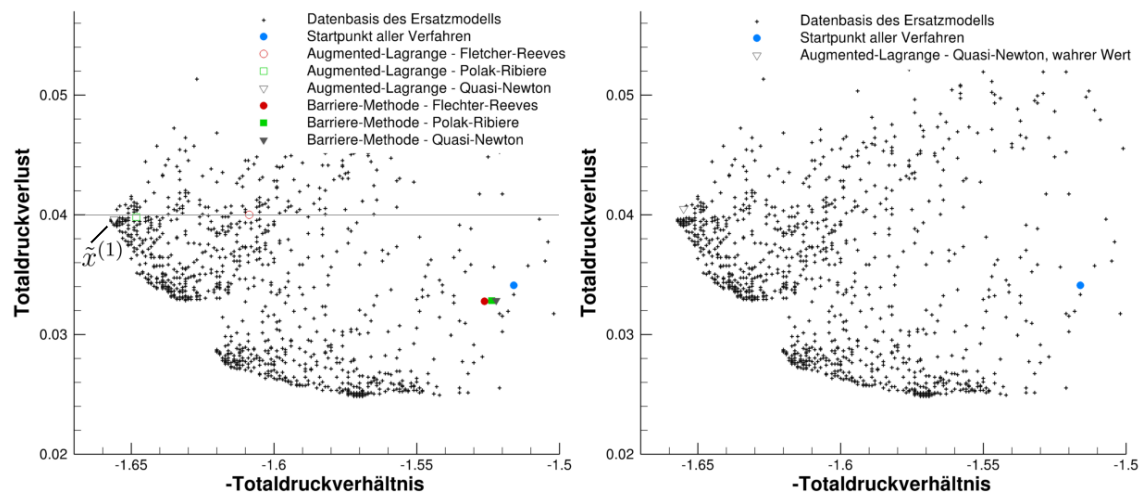


Abbildung 5.27: Zielfunktionswerte der Ersatzmodell-Datenbasis in schwarz, Approximation der Zielfunktionswerte aller Endpunkte der konvergierten Verfahren (links) sowie wahre Werte des Augmented-Lagrange-/Quasi-Newton-Verfahrens (rechts) bei Restriktion des Totaldruckverlustes durch 0.04.

punkt $\tilde{x}^{(1)} = (100.08, 137.83, 90.00, 0.28, 0.78)^T$ des Augmented-Lagrange-/Quasi-Newton-Verfahrens mit einem approximierten Totaldruckverhältnis von 1.656 und einem approximierten Totaldruckverlust von 0.039 interessant.

Die unterschiedlichen Verhaltensweisen der Barriere-Methode und des Augmented-Lagrange-Verfahrens bei Auftreten einer Nebenbedingung sind in Abbildung 5.27, links, erkennbar. Die Barriere-Methode konvergiert als Innere-Punkte-Methode mit allen Verfahren zur unrestringierten Minimierung gegen lokale Minima mit einem Totaldruckverlust von 0.0328. Dieser liegt somit weit unter der Grenze von 0.04. Das Augmented-Lagrange-Verfahren konvergiert hingegen gegen lokale Minima mit einem Totaldruckverlust nahe der zulässigen oberen Grenze.

Da wir in diesem Fall den Totaldruckverlust nicht minimieren, sondern restringieren, liegt das Augenmerk auf dem zu maximierenden Totaldruckverhältnis. Hier konvergiert das Augmented-Lagrange-Verfahren gegen die weitaus besseren Punkte (vgl. Abb. 5.27, links), jedoch sei erneut darauf hingewiesen, dass die Verfahren gegen lokale und nicht globale Minima konvergieren.

Nach Abschluss der Optimierung des Ersatzmodells berechnen wir die wahren Funktionswerte für den vielversprechenden Punkt $\tilde{x}^{(1)}$ des Augmented-Lagrange-/Quasi-Newton-Verfahrens mit MISES. Wir erhalten ein Totaldruckverhältnis von 1.655 und einen Totaldruckverlust von 0.0405 im Vergleich zu den approximierten Werten 1.656 respektive 0.039. Diese gute Übereinstimmung der vorhergesagten und wahren Werte spiegelt sich erstmals auch in den Unsicherheiten wieder: Das Direct GEK berechnet für das approximierte Totaldruckverhältnis eine geringe Unsicherheit von $h_1(\tilde{x}^{(1)}) = 0.044$ und für den Totaldruckverlust die geringe Unsicherheit $h_2(\tilde{x}^{(1)}) = 0.029$.

Erstmals erreichen unsere Verfahren nicht nur im Ersatzmodell ein gutes Optimierungsergebnis, sondern auch bei der darauffolgenden Überprüfung mit MISES. Der berechnete Punkt erreicht zwar nicht exakt die Werte -1.657 und 0.0395, -1.656 und 0.0392 oder -1.655 und 0.0390 der zuvor mit AutoOpti berechneten umliegenden Pareto-Optima, allerdings wurde $\tilde{x}^{(1)}$ auch mit Hilfe des Ersatzmodells und nicht direkt mit MISES berechnet. Vergleichen wir die Einträge dieser eben genannten optimalen Verdichterprofilkonfigurationen mit den Einträgen von $\tilde{x}^{(1)}$, so ist eine gute Übereinstimmung der Werte aller fünf freien Variablen erkennbar. Der Punkt $\tilde{x}^{(1)}$ befindet sich folglich tatsächlich nahe den umliegenden Pareto-Optima.

Restriktion des Totaldruckverlustes durch 0.03

In diesem Abschnitt wird die obere Grenze des Totaldruckverlustes auf den Wert 0.03 herabgesetzt. Das Totaldruckverhältnis soll weiterhin maximiert werden. Da der im vorherigen Abschnitt verwendete Startpunkt wiederum die nun stärkere Nebenbedingung an den Totaldruckverlust verletzt, betrachten wir im Folgenden den neuen Startpunkt $x^{(0)} = (107.918, 138.414, 90.710, 0.297, 0.737)^T$ (vgl. Abb. 5.28).

In Abbildung 5.28, links, wird das unterschiedliche Verhalten der beiden restringierten Verfahren erneut sichtbar. Die Endpunkte des Augmented-Lagrange-Verfahrens befinden sich wiederholt nahe der oberen Grenze, während die Barriere-Methode gegen Punkte konvergiert, deren Totaldruckverlust deutlich von dieser entfernt ist.

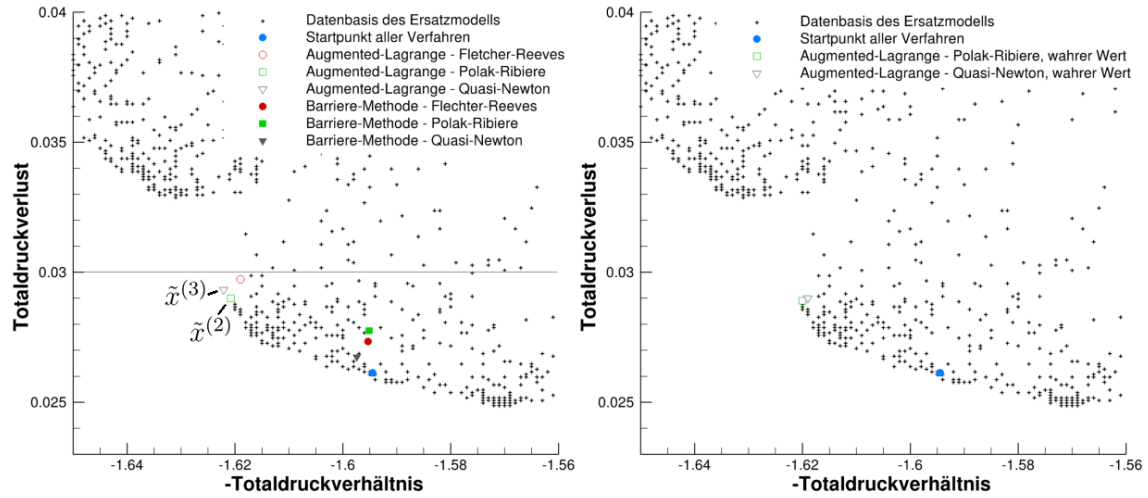


Abbildung 5.28: Zielfunktionswerte der Ersatzmodell-Datenbasis in schwarz, Approximation der Zielfunktionswerte aller Endpunkte der konvergierten Verfahren (links) sowie wahre Werte ausgewählter Verfahren (rechts) bei Restriktion des Totaldruckverlustes durch 0.03.

Besonders vielversprechend sind die beiden Endpunkte

$$\begin{aligned}\tilde{x}^{(2)} &= (107.92, 138.41, 90.71, 0.30, 0.74)^T \text{ mit } p_{\text{TPR}} = 1.621 \text{ und } p_{\text{TPL}} = 0.0290 \text{ und} \\ \tilde{x}^{(3)} &= (108.42, 137.16, 90.00, 0.30, 0.78)^T \text{ mit } p_{\text{TPR}} = 1.622 \text{ und } p_{\text{TPL}} = 0.0293\end{aligned}$$

des Augmented-Lagrange-Verfahrens mit dem Polak-Ribière- respektive Quasi-Newton-Verfahren zur unrestringierten Minimierung.

Nach Abschluss der Optimierung berechnet MISES die wahren Funktionswerte der beiden eben genannten Punkte. Für $\tilde{x}^{(2)}$ erhalten wir das Totaldruckverhältnis 1.620 und den Totaldruckverlust 0.0289, für $\tilde{x}^{(3)}$ die Werte 1.619 und 0.0290 (vgl. Abb. 5.28, rechts). Die sehr gute Übereinstimmung der Vorhersagen mit den wahren Werten spiegelt sich erneut in den Unsicherheiten wieder, welche eine ähnlich geringe Größenordnung wie im vorherigen Abschnitt aufweisen.

Somit ist es uns erneut gelungen, Punkte beziehungsweise Verdichterprofilkonfigurationen zu ermitteln, welche nicht nur bei Betrachtung des Ersatzmodells, sondern insbesondere bei Untersuchung der wahren Zielfunktionswerte nahe den durch AutoOpti berechneten Pareto-Optima liegen.

Restriktion des Totaldruckverlustes durch 0.025

Im folgenden Abschnitt soll der Totaldruckverlust noch stärker restringiert werden. Wir wählen für diesen als neue obere Grenze den Wert 0.025. Wie Abbildung 5.29, links, verdeutlicht, existieren nur wenige Punkte in der Datenbasis, welche diese Nebenbedingung erfüllen. Der Punkt $x^{(0)} = (107.520, 138.725, 91.903, 0.299, 0.703)^T$ sei der neue Startpunkt der Verfahren (vgl. Abb. 5.29).

In diesem Abschnitt untersuchen wir eine sehr starke Restriktion des Totaldruckverlustes, um eventuelle Grenzen der Verfahren aufzuzeigen und falls möglich Verdichterprofilkonfigurationen in der Pareto-Front zu errechnen.

Wie der in Abbildung 5.29, links, eingezeichnete Endpunkt des Fletcher-Reeves-Verfahrens veranschaulicht, hat die Barriere-Methode nun größere Schwierigkeiten gegen Punkte zu konvergieren, deren Totaldruckverlust erkennbar von der oberen Grenze entfernt ist. Auch die beiden verbleibenden Endpunkte der Barriere-Methode besitzen einen Totaldruckverlust, der im Vergleich zu den Verlusten der vorherigen Abschnitte deutlich näher an der oberen Grenze liegt.

Im Anschluss an die Optimierung berechnen wir die wahren Funktionswerte der vier für uns interessanten Punkte (vgl. Abb. 5.29, links)

$$\begin{aligned}\tilde{x}^{(4)} &= (107.87, 137.97, 90.03, 0.30, 0.72)^T \text{ mit } p_{\text{TPR}} = 1.590 \text{ und } p_{\text{TPL}} = 0.0250, \\ \tilde{x}^{(5)} &= (107.82, 137.87, 90.00, 0.30, 0.73)^T \text{ mit } p_{\text{TPR}} = 1.595 \text{ und } p_{\text{TPL}} = 0.0250, \\ \tilde{x}^{(6)} &= (107.00, 139.27, 91.83, 0.29, 0.71)^T \text{ mit } p_{\text{TPR}} = 1.578 \text{ und } p_{\text{TPL}} = 0.0247 \text{ und} \\ \tilde{x}^{(7)} &= (107.26, 139.31, 92.48, 0.29, 0.71)^T \text{ mit } p_{\text{TPR}} = 1.573 \text{ und } p_{\text{TPL}} = 0.0247\end{aligned}$$

mit Hilfe von MISES. Abbildung 5.29, rechts, zeigt die Resultate des Strömungslösers. Wir erhalten für die Punkte $\tilde{x}^{(4)}$ bis $\tilde{x}^{(7)}$ jeweils das folgende wahre Totaldruckverhältnis und den wahren Totaldruckverlust 1.590 und 0.0259, 1.594 und 0.0261, 1.580 und 0.0254 respektive 1.574 und 0.0253. Wie aufgrund der Abbildung 5.29, links, zu erwarten ist, werden die Totaldruckverluste für alle betrachteten Punkte zu gut vorhergesagt. Die Vorhersagen des Totaldruckverhältnisses sind hingegen richtig. Dieser Unterschied zwischen der guten Vorhersage des Totaldruckverhältnisses und der relativ ungenauen Vorhersage des Totaldruckverlustes spiegelt sich nicht in den Unsicherheiten wider. Hier erhalten wir für alle betrachteten Werte geringe Unsicherheiten zwischen 0.01 und 0.05.

Anhand dieser starken Restriktion des Totaldruckverlustes konnten wir nachweisen, dass die Verfahren weiter stabil gegen lokale Minima konvergieren und dabei die Nebenbedingung nicht verletzen.

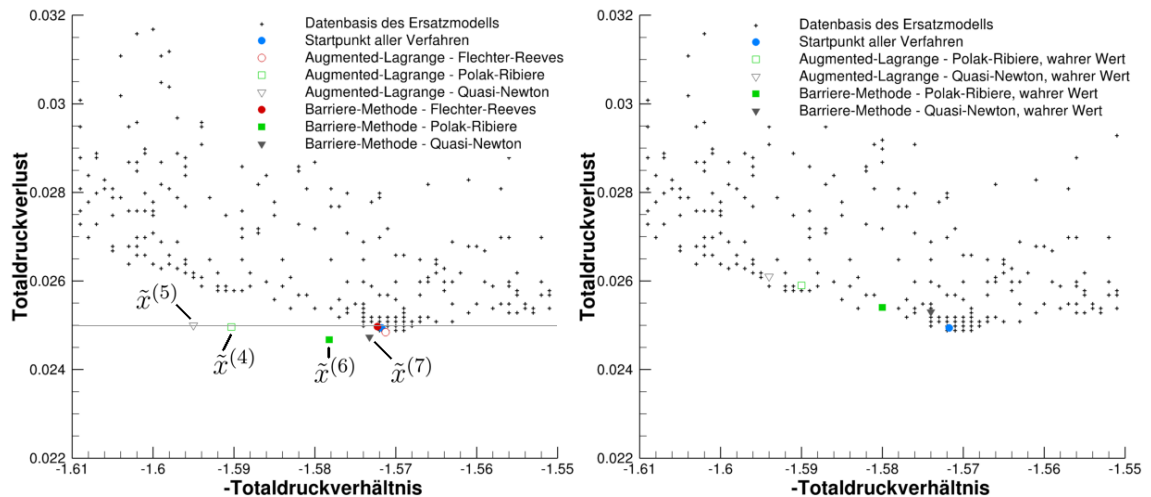


Abbildung 5.29: Zielfunktionswerte der Ersatzmodell-Datenbasis in schwarz, Approximation der Zielfunktionswerte aller Endpunkte der konvergierten Verfahren (links) sowie wahre Werte ausgewählter Verfahren (rechts) bei Restriktion des Totaldruckverlustes durch 0.025.

Restriktion des Totaldruckverlustes durch 0.0325

Im folgenden Abschnitt untersuchen wir, inwieweit es möglich ist, die Lücke in der Pareto-Front bei einem Totaldruckverhältnis von 1.62 (vgl. Abb. 5.20) zu schließen. Das zuvor untersuchte Problem mit Restriktion des Totaldruckverlustes durch den Wert 0.03 (vgl. S. 81) gab aufgrund einer zu starken Restriktion keine zufriedenstellende Antwort bezüglich dieser Frage. Aus diesem Grund begrenzen wir im Folgenden den Totaldruckverlust durch 0.0325 nach oben. Der Punkt $x^{(0)} = (107.87, 138.56, 90.51, 0.30, 0.78)^T$ sei der neue Startpunkt aller Verfahren.

Abbildung 5.30 zeigt die approximierten Funktionswerte der Endpunkte aller untersuchten Verfahren. Selbst bei Lockerung der Restriktion gelingt es keinem der Verfahren die Lücke der Pareto-Front zu schließen. Die Verfahren konvergieren erneut gegen Punkte mit approximierten Zielfunktionswerten nahe den Endpunkten bei Restriktion durch 0.03 (vgl. Abb. 5.28, links).

Die Lockerung der Restriktion des Totaldruckverlustes hat keine entsprechende Erhöhung von diesem zur Folge. Anscheinend existieren keine Verdichterprofilkonfigurationen, die in diesem Bereich des Totaldruckverhältnisses und Totaldruckverlustes konvergieren.

Das vorangegangene Unterkapitel zeigt, dass die Maximierung des Totaldruckverhältnisses bei Restriktion des Totaldruckverlustes deutlich bessere Resultate als die in Abschnitt 5.2.1 betrachtete Optimierung der gewichteten Summe liefert. Die bessere Kontrolle der möglichen Endpunkte konvergierter Verfahren stellt einen bemerkenswerten Vorteil dieser Herangehensweise dar.

Des Weiteren konnten wir in diesem Abschnitt ein unterschiedliches Verhalten der beiden restringierten Verfahren im Umgang mit der Nebenbedingung an den Totaldruckverlust

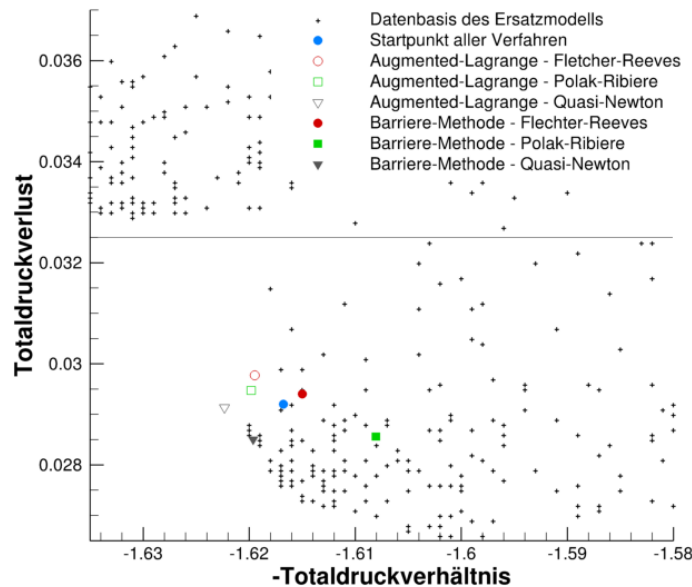


Abbildung 5.30: Zielfunktionswerte der Ersatzmodell-Datenbasis in schwarz und Approximation der Zielfunktionswerte aller Endpunkte der konvergierten Verfahren bei Restriktion des Totaldruckverlustes durch 0.0325.

registrieren. Die Barriere-Methode konvergiert als Innere-Punkte-Methode vornehmlich gegen deutlich von der oberen Grenze entfernte Punkte. Die Endpunkte des Augmented-Lagrange-Verfahrens können hingegen die Grenzen des Totaldruckverlustes und der freien Variablen tangieren.

Diese unterschiedlichen Verhaltensweisen müssen bei einer Optimierung entsprechend berücksichtigt werden. Um gleiche Ergebnisse zu erhalten, könnte man die Restriktion der Barriere-Methode gegenüber der des Augmented-Lagrange-Verfahrens lockern.

Die verschiedenen Konvergenzverläufe der Verfahren wurden im vorherigen Abschnitt nicht weiter betrachtet, da das Augenmerk dieses und der nachfolgenden Abschnitte auf der Untersuchung verschiedener Herangehensweisen zur Lösung des Mehrziel-Verdichterproblems liegt.

Grundsätzlich zeigen die Verfahren das bereits bekannte Konvergenzverhalten. Insbesondere weisen die restringierten Verfahren mit dem Quasi-Newton-Verfahren zur unrestringierten Minimierung weiterhin eine vergleichbar geringe Iterationsanzahl auf.

5.2.3 Optimierung der gewichteten Summe mit Restriktion

Durch gewichtete Addition der Zielfunktionen *Totaldruckverhältnis* und *Totaldruckverlust* wird das Mehrzieloptimierungsproblem erneut auf ein Problem der Form (P) reduziert. Hierbei verwenden wir die schon untersuchte Gewichtung durch das arithmetische Mittel der Funktionswerte aller 1206 in der Datenbasis enthaltenen Verdichterprofile. Somit wird das Totaldruckverhältnis mit dem Faktor $-\frac{1}{1.593} \approx -0.628$ und der Totaldruckverlust mit dem Faktor $\frac{1}{0.037} \approx 27.027$ gewichtet.

Im folgenden Abschnitt betrachten wir verschieden starke Restriktionen der Unsicherheiten $h_1(x)$ und $h_2(x)$ des Totaldruckverhältnisses respektive Totaldruckverlustes. Die freien Variablen werden wie in (5.1) gefordert restringiert.

Der Punkt $x^{(0)} = (107.72, 137.94, 90.24, 0.30, 0.77)^T$ sei Startpunkt aller Verfahren. Dieser weist aufgrund seiner Lage nahe der Pareto-Front mit vielen beprobten Punkten in der Nähe (siehe bspw. Abb. 5.31) die vergleichbar geringen Unsicherheiten $h_1(x^{(0)}) = 0.035$ und $h_2(x^{(0)}) = 0.018$ auf. Solch geringe Unsicherheiten des Startpunktes sind notwendig, um in den folgenden Abschnitten die Auswirkungen bei Änderung der Unsicherheitsrestriktionen untersuchen zu können.

Optimierung bei Restriktion $h_1(x) \leq 0.04$ und $h_2(x) \leq 0.02$

Der folgende Abschnitt beschreibt die Auswirkungen bei Restriktion der Unsicherheiten des Totaldruckverhältnisses und Totaldruckverlustes durch 0.04 respektive 0.02 nach oben. Diese sehr starke Restriktion schränkt die Menge der zulässigen Punkte ein. Selbst der Startpunkt $x^{(0)}$ mit den relativ geringen Unsicherheiten 0.035 beziehungsweise 0.018 genügt der Restriktion nur relativ knapp.

Abbildung 5.31, links, zeigt die approximierten Zielfunktionswerte der Endpunkte aller untersuchten Verfahren. Die starke Restriktion hat für die Endpunkte der Barriere-Methode zur Folge, dass diese sich kaum vom Startpunkt entfernen. Ein Vergleich der linken und rechten Abbildung 5.31 zeigt jedoch auch, dass die vorhergesagten und wahren Zielfunktionswerte der Barriere-Methode keine große Differenz aufweisen. Aufgrund der starken Unsicherheitsrestriktion stimmt die Vorhersage des Direct GEK gut mit den wahren Zielfunktionswerten überein.

Im Vergleich zur Barriere-Methode zeigen die Endpunkte des Augmented-Lagrange-Verfahrens ein anderes Verhalten. Aufgrund der Abbruchbedingung $n(x^{(k+1)}) \leq \eta_k$ in Schritt 3 des Augmented-Lagrange-Verfahrens 4.2 besteht die Möglichkeit, dass die Endpunkte des Verfahrens eine geringe Unzulässigkeit aufweisen. Hierbei ist $n = n(x)$ die Nebenbedingungsverletzung (vgl. Formel (4.2), S. 37).

Da die auftretenden freien Variablen und die Unsicherheiten h_1 sowie h_2 des Totaldruckverhältnisses respektive Totaldruckverlustes nicht normiert betrachtet werden, hat eine geringe Verletzung der relativ kleinen Unsicherheitsrestriktionen kaum Gewicht. Folglich besitzen die Augmented-Lagrange-Endpunkte trotz Unsicherheitsrestriktion durch 0.04 beziehungsweise 0.02 die Unsicherheiten $h_1(x) \approx 0.05$ und $h_2(x) \approx 0.029$ und sind somit unzulässig.

Die geringfügige Überschreitung der oberen Grenze 0.3 der freien Variable y_{SS} durch den im Endpunkt angenommenen Wert 0.302 kann ebenso mit Hilfe obiger Erläuterung begründet werden: Diese Überschreitung der oberen Grenze hat eine noch kleinere Nebenbedingungsverletzung als die der Unsicherheiten h_1 respektive h_2 zur Folge und fällt somit nicht ins Gewicht.

Ist die strikte Zulässigkeit der Punkte gefordert, so müssen die Restriktionen an die freien Variablen und Unsicherheiten entsprechend verstärkt werden.

Vergleichen wir die approximierten und wahren Zielfunktionswerte der Augmented-Lagrange-Endpunkte in Abbildung 5.31, so können wir eine gute Vorhersage des Totaldruckverhältnisses und relativ große Differenzen beim Totaldruckverlust feststellen. Obwohl die Unsicherheitsrestriktion durch das Augmented-Lagrange-Verfahren überschritten wurde, kann eine solche Abweichung der vorhergesagten und wahren Totaldruckverlustwerte nicht durch die weiterhin relativ geringe Unsicherheit $h_2(x) \approx 0.029$ erklärt werden. Die durch das Direct GEK berechnete Unsicherheit ist in diesem Fall erneut zu gering.

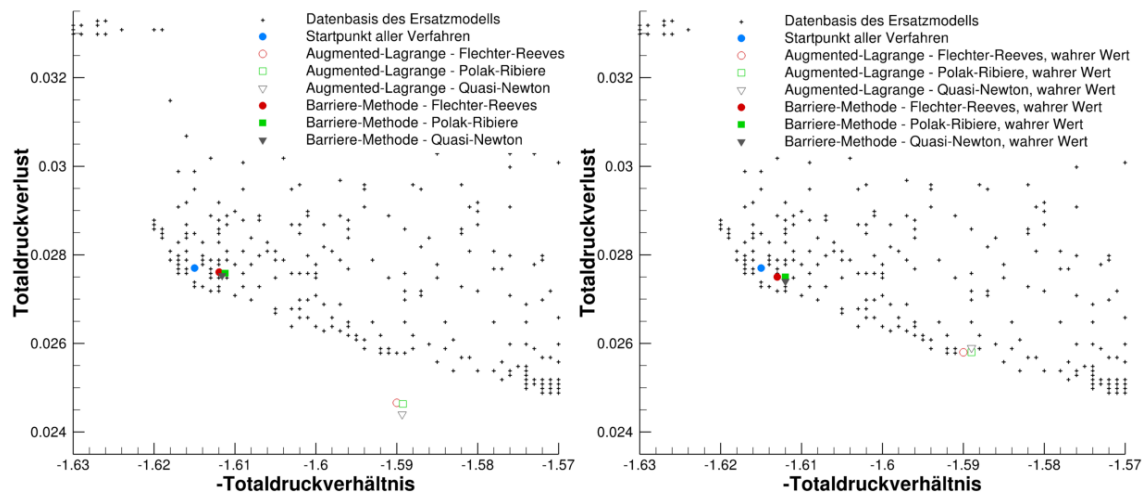


Abbildung 5.31: Zielfunktionswerte der Ersatzmodell-Datenbasis in schwarz, Approximation der Zielfunktionswerte aller Endpunkte der konvergierten Verfahren (links) sowie wahre Werte ausgewählter Verfahren (rechts) bei Restriktion der Unsicherheiten durch $h_1(x) \leq 0.04$ und $h_2(x) \leq 0.02$.

Optimierung bei schwächeren Restriktionen

Im folgenden Abschnitt analysieren wir die Auswirkungen bei Lockerung der Unsicherheitsrestriktionen. Die drei Fälle

$$\begin{aligned} h_1(x) &\leq 0.05 \text{ und } h_2(x) \leq 0.025, \\ h_1(x) &\leq 0.06 \text{ und } h_2(x) \leq 0.03 \text{ sowie} \\ h_1(x) &\leq 0.09 \text{ und } h_2(x) \leq 0.045 \end{aligned}$$

werden betrachtet. Abbildung 5.33 zeigt die approximierten und wahren Zielfunktionswerte aller konvergierten Verfahren der drei obigen Fälle im Vergleich.

Im Verlauf der fortschreitenden Lockerung der Unsicherheitsrestriktionen (in Abb. 5.33 von oben nach unten) entfernen sich die approximierten Zielfunktionswerte der Barriere-Methode weiter vom Startpunkt (vgl. Abb. 5.33, links). Den Verfahren stehen nun mehr zulässige Punkte für die Optimierung zur Verfügung. Aufgrund der relativ starken Gewichtung des Totaldruckverlustes konvergieren die Verfahren gegen Punkte mit einem geringeren Totaldruckverlust und folglich einem geringeren Totaldruckverhältnis. Gleiches Verhalten konnten wir schon in Abschnitt 5.2.1 bei Gewichtung durch die Faktoren -0.628 und 27.027 feststellen.

Ein Vergleich der approximierten und wahren Funktionswerte zeigt eine Folge der schwächeren Restriktionen. Die Barriere-Methode weist bei fortschreitender Lockerung grundsätzlich eine höhere Differenz der approximierten und wahren Zielfunktionswerte auf. In Abbildung 5.32 ist die Norm des Differenzvektors

$$\begin{pmatrix} p_{\text{TPR, approx}} - p_{\text{TPR, wahr}} \\ p_{\text{TPL, approx}} - p_{\text{TPL, wahr}} \end{pmatrix}$$

für die verschiedenen, untersuchten Unsicherheitsrestriktionen und Verfahren angegeben. Trotz Schwankungen erkennen wir überwiegend einen Aufwärtstrend der Norm des Differenzvektors für alle unrestringierten Verfahren der Barriere-Methode. Wie zu erwarten werden die Zielfunktionswerte bei schwächerer Restriktion meist schlechter vorhergesagt. Die approximierten Zielfunktionswerte des Augmented-Lagrange-Verfahrens weisen keine Änderung bei Lockerung der Unsicherheitsrestriktionen auf (siehe Abb. 5.33, links, von oben nach unten). Wie zuvor bei Restriktion durch 0.04 respektive 0.02 konvergieren die Verfahren unabhängig von den gegenwärtigen Unsicherheitsrestriktionen gegen Punkte mit einem approximierten Totaldruckverhältnis von etwa 1.59 und einem Totaldruckverlust von circa 0.0245.

In den beiden Fällen $h_1(x) \leq 0.05$ und $h_2(x) \leq 0.025$ sowie $h_1(x) \leq 0.06$ und $h_2(x) \leq 0.03$ treten erneut geringe Verletzungen der Unsicherheitsrestriktionen auf. Aufgrund der schwachen Restriktionen im letzten Fall werden die Nebenbedingungen an die Unsicherheiten hier nicht weiter verletzt.

In einigen Endpunkten des Augmented-Lagrange-Verfahrens tritt weiterhin eine geringfügige Überschreitung der oberen Grenze der Variable y_{SS} auf. Die Verletzungen der Nebenbedingungen lassen sich wie im vorherigen Abschnitt begründen.

Erneut wird das Totaldruckverhältnis der Augmented-Lagrange-Endpunkte gut und der Totaldruckverlust fehlerhaft vorhergesagt (vgl. Abb. 5.33). Die Unabhängigkeit der approximierten Zielfunktionswerte von der Unsicherheitsrestriktion spiegelt sich in der Norm des Differenzvektors wieder: Die in Abbildung 5.32 verzeichneten Verläufe für die un-

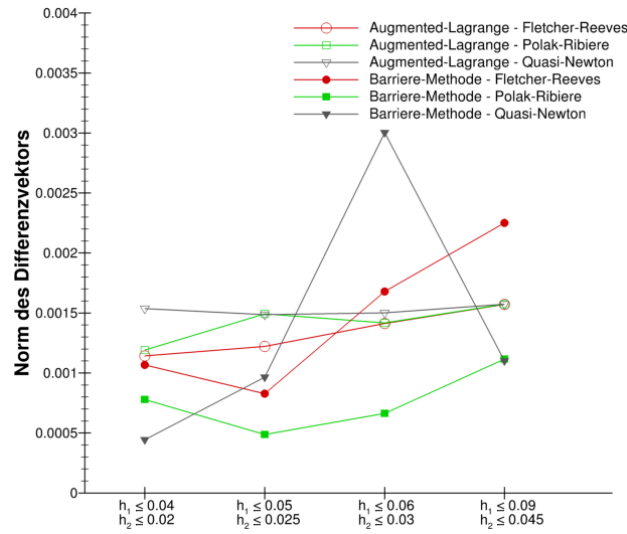


Abbildung 5.32: Norm $\|(p_{\text{TPR, approx}} - p_{\text{TPR, wahr}}, p_{\text{TPL, approx}} - p_{\text{TPL, wahr}})^T\|$ des Differenzvektors bei unterschiedlicher Restriktion der Unsicherheiten für alle untersuchten Verfahren

tersuchten unrestringierten Verfahren beim Augmented-Lagrange-Verfahren zeigen kaum Veränderung bei Lockerung der Restriktionen.

Betrachten wir Abbildung 5.33, unten, so erkennen wir, dass die Endpunkte der Barriere-Methode und des Augmented-Lagrange-Verfahrens aufgrund der Lockerung der Unsicherheitsrestriktionen ähnliche Totaldruckverhältnisse aufweisen. Eine Differenz der Funktionswerte besteht jedoch sowohl beim approximierten als auch beim wahren Totaldruckverlust. Diese Differenz entsteht aufgrund der Inneren-Punkte-Eigenschaft der Barriere-Methode. Im Gegensatz zum Augmented-Lagrange-Verfahren, dessen Endpunkte eine geringe Verletzung der Nebenbedingung an y_{SS} aufweisen, ist dies bei der Barriere-Methode nicht möglich. Die Endpunkte können hier nur gegen die obere Grenze konvergieren, erreichen diese jedoch niemals. Folglich sind die Zielfunktionswerte der Barriere-Methode in diesem Fall schlechter als die des teilweise unzulässigen Augmented-Lagrange-Verfahrens.

Zusammenfassend zeigt die Barriere-Methode bei unterschiedlich starker Restriktion der Unsicherheiten das zu erwartende Verhalten. Insbesondere kann die zuvor in Abschnitt 5.1.4, Seite 69, beschriebene Vorgehensweise anhand der Barriere-Methode nachvollzogen werden: Eine zu schwache Restriktion der Unsicherheit, wie in den beiden letzten Fällen $h_1(x) \leq 0.09$ respektive $h_2(x) \leq 0.045$ sowie $h_1(x) \leq 0.06$ und $h_2(x) \leq 0.03$ führt zu einer teilweise guten Vorhersage der Zielfunktionswerte durch das Direct GEK. Dies zeigen beispielsweise die Endpunkte der Barriere-Methode mit Polak-Ribière-Verfahren bei Restriktion durch 0.09 respektive 0.045 oder mit Fletcher-Reeves- beziehungsweise Quasi-Newton-Verfahren bei Restriktion durch 0.06 respektive 0.03 (vgl. Abb. 5.33, links). Die mit Hilfe von MISES berechneten wahren Funktionswerte unterscheiden sich in diesen Fällen jedoch signifikant von der Vorhersage, was die relativ große Norm des Differenzvektors in Abbildung 5.32 für die Barriere-Methode mit Polak-Ribière-Verfahren bei Restriktion durch 0.09 respektive 0.045 und mit Fletcher-Reeves- beziehungsweise Quasi-Newton-Verfahren bei Restriktion durch 0.06 respektive 0.03 unterstreicht.

Aufgrund einer Verstärkung der Restriktionen durch $h_1(x) \leq 0.05$ respektive $h_2(x) \leq 0.025$ werden die Funktionswerte zwar teilweise mit schlechteren Zielfunktionswerten vorhergesagt (siehe Polak-Ribière-Verfahren), dafür stimmen die vorhergesagten und wahren Werte gut überein (vgl. Abb. 5.32) und bilden einen akzeptablen Kompromiss zwischen Stärke der Restriktion und guten Zielfunktionswerten.

Eine wie im vorherigen Abschnitt vorgenommene weitere Verstärkung der Restriktionen führt zu einer Stagnation der Verfahren in der Umgebung des Startpunktes mit geringen Unsicherheiten, jedoch relativ schlechten Zielfunktionswerten.

Die Überprüfung dieser Vorgehensweise ist mit dem in Abschnitt 5.2.1 verwendeten Startpunkt $x^{(0)} = (114.9, 136, 100, 0.2170, 0.7447)^T$ nicht möglich, da die Startpunkte der Barriere-Methode zulässig sein müssen und die Unsicherheit $h_1(x^{(0)}) = 0.180$ des Totaldruckverhältnisses folglich einen zu großen Wert für Untersuchungen mit unterschiedlich starken Restriktionen aufweist.

Beim Augmented-Lagrange-Verfahren konnten wir aufgrund der zu kleinen Wertebereiche der Unsicherheitsrestriktionen keine Veränderung bei Betrachtung verschiedener Restriktionen feststellen. Dies liegt in der Abbruchbedingung des Augmented-Lagrange-Verfahrens begründet, welche eine geringe Nebenbedingungsverletzung zulässt. Ist die strikte Zulässigkeit gefordert, so kann eine Normierung aller freien Variablen, Zielfunktionen und Nebenbedingungen, in diesem Fall der Unsicherheiten h_1 und h_2 , und eine Verstärkung der Restriktionen hilfreich sein. Eine andere Möglichkeit besteht in der Verwendung der strikt zulässigen Barriere-Methode.

5.2.4 Zusammenfassung der Verdichterdesign-Ergebnisse

Die vorangegangenen Auswertungen zeigen, dass die Optimierung des Verdichterdesignbeispiels mit Hilfe der im Rahmen dieser Arbeit implementierten Verfahren möglich ist. Einzig das Gradientenverfahren weist bei beiden Verfahren zur restringierten Optimierung das bekannte, langsame Konvergenzverhalten auf und wurde aus diesem Grund nicht weiter untersucht.

Die in Abschnitt 5.2.2 aufgezeigte Herangehensweise mit Optimierung des Totaldruckverhältnisses unter Restriktion des Totaldruckverlustes zeigt die besten Resultate. Die approximierten und wahren Zielfunktionswerte stimmen gut überein und viele der nachgerechneten Punkte liegen nahe der Pareto-Front. Des Weiteren ist der Bereich der Zielfunktionswerte, gegen welche die Endpunkte konvergieren, durch den Benutzer aufgrund der Restriktion des Totaldruckverlustes besser eingrenzbar.

Abschnitt 5.2.1 zeigt, dass unterschiedliche Gewichtungen der beiden Zielfunktionen zwar grundsätzliche Tendenzen in den Funktionswerten der Endpunkte erkennen lassen, diese Herangehensweise jedoch schwer steuerbar ist. In diesem Fall kann der Benutzer versuchen, durch sukzessives Ändern der Gewichtungsfaktoren die gewünschten Wertebereiche der Zielfunktionen zu erreichen. Aufgrund der fehlerhaften Vorhersage des Totaldruckverlustes durch das Direct GEK wiesen die approximierten Zielfunktionswerte teilweise sehr gute Werte auf, welche jedoch nach anschließender Berechnung mit MISES revidiert wurden. Hier ist eine Restriktion der Unsicherheit notwendig. Diese wurde in Abschnitt 5.2.3 untersucht, allerdings aus auf Seite 89 erläuterten Gründen mit einem veränderten Startpunkt.

Des Weiteren wurden in Abschnitt 5.2.3 die Auswirkungen unterschiedlich starker Restriktionen der Unsicherheiten des Totaldruckverhältnisses und Totaldruckverlustes untersucht. Für die strikt zulässige Barriere-Methode konnten wir die in Abschnitt 5.1.4 vorgeschlagene Herangehensweise nachvollziehen und einen guten Kompromiss zwischen Restriktionen der Unsicherheiten und Optimierung der Zielfunktionen finden. Die Unsicherheitsrestriktion bei der Augmented-Lagrange-Methode schlug aufgrund der Abbruchbedingung fehl. Die mit MISES berechneten Zielfunktionswerte der Endpunkte des Augmented-Lagrange-Verfahrens befinden sich teilweise nahe der Pareto-Front, weisen allerdings in einigen Fällen aufgrund der Abbruchbedingung Unzulässigkeiten auf.

Zusätzlich zu den bisher betrachteten Optimierungsproblemen wurde eine weitere Optimierung mit einer auf 230 Verdichterprofilkonfigurationen ausgedünnten Datenbasis für das Direct GEK vorgenommen. Wie in Abschnitt 5.2.1 wurde derselbe Startpunkt $x^{(0)} = (114.9, 136, 100, 0.2170, 0.7447)^T$ verwendet. Erneut wurden unterschiedlich stark gewichtete Summen der beiden Zielfunktionen betrachtet.

Die Verfahren zeigen ein ähnliches Verhalten wie in Abschnitt 5.2.1. Das Totaldruckverhältnis wird gut approximiert, der Totaldruckverlust hingegen fehlerhaft. Dieses Optimierungsproblem wird im Folgenden nicht näher betrachtet, da keine neuen Erkenntnisse gewonnen werden konnten.

Insgesamt haben wir das Ziel erreicht und durch verschiedene Herangehensweisen Möglichkeiten aufgezeigt, Punkte beziehungsweise Verdichterprofilkonfigurationen nahe der Pareto-Front zu konstruieren. Da die zuvor mit AutoOpti durchgeführte Optimierung und Erzeugung der Datenbasis schon sehr gute Pareto-Optima lieferte, ist es uns nicht gelungen, neue Pareto-Optima zu berechnen.

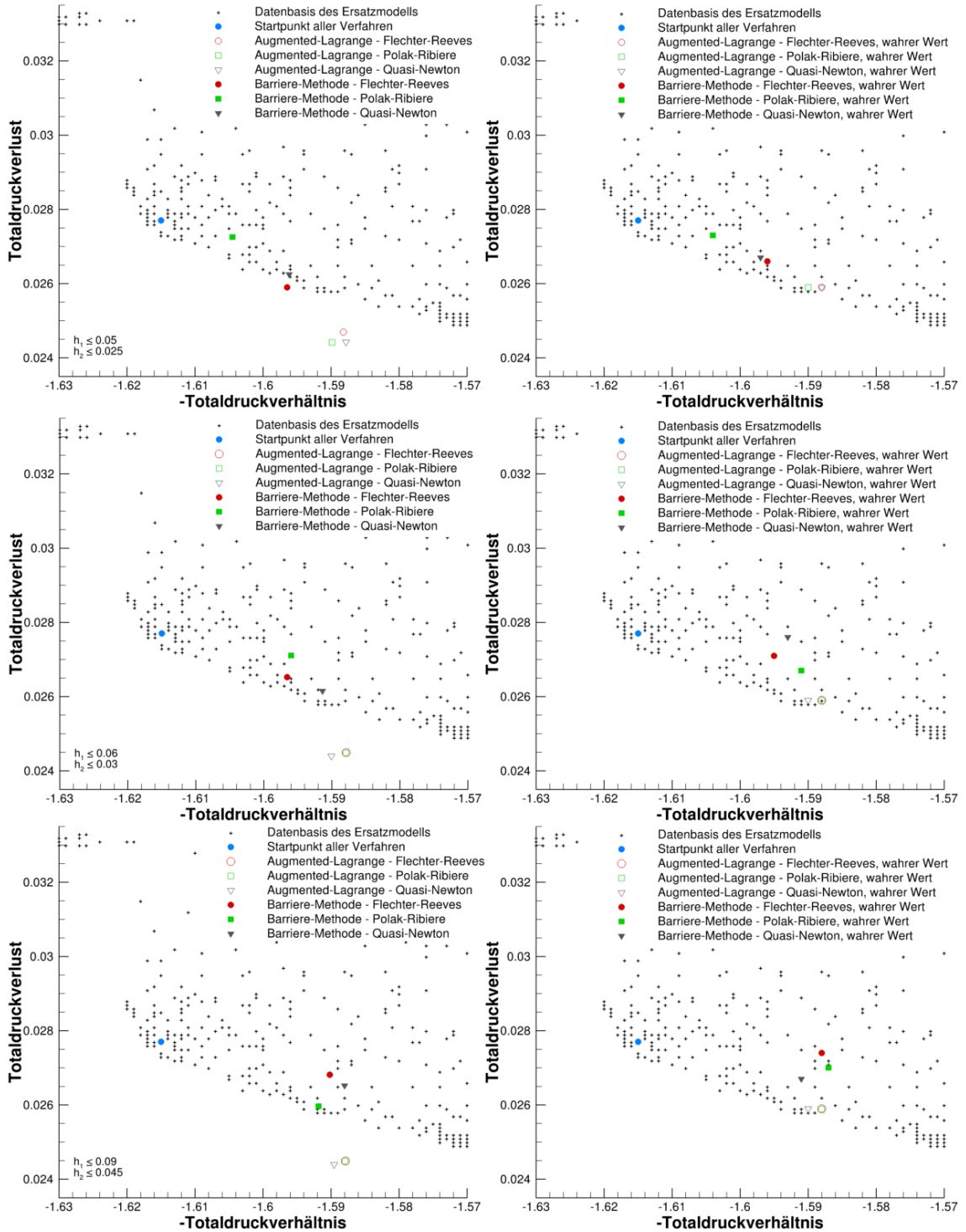


Abbildung 5.33: Zielfunktionswerte der Ersatzmodell-Datenbasis in schwarz, Approximation der Zielfunktionswerte aller Endpunkte der konvergierten Verfahren (links) sowie wahre Werte ausgewählter Verfahren (rechts) bei Restriktion der Unsicherheiten durch $h_1(x) \leq 0.05$ und $h_2(x) \leq 0.025$ (oben), durch $h_1(x) \leq 0.06$ und $h_2(x) \leq 0.03$ (Mitte) und durch $h_1(x) \leq 0.09$ und $h_2(x) \leq 0.045$ (unten).

6 Fazit und Ausblick

Zu Beginn dieser Arbeit haben wir einen Einblick in den bisherigen Stand der Verdichterprofiloptimierung am Institut für Antriebstechnik des DLR gegeben. Das dort verwendete, auf genetischen Algorithmen basierende Optimierungsprogramm AutoOpti wurde vorgestellt und die Gründe für eine Gradienten-basierte Optimierung des Ersatzmodells dargelegt. Die im Rahmen dieser Arbeit implementierten Gradienten-basierten Optimierungsalgorithmen sollen als Teil des Verdichterdesignprozesses in AutoOpti zum Einsatz kommen.

Im Anschluss wurden zum besseren Verständnis die mathematischen Hintergründe verschiedener Kriging-Methoden erläutert, mit deren Hilfe die zur Optimierung benötigten Ersatzmodelle erstellt werden. Da wir zu diesen geostatistischen Interpolationsverfahren keine passende exakte mathematische Literatur finden konnten, beruhen die Beweise in diesem Kapitel auf eigenen Überlegungen.

Daraufhin wurden die im Rahmen dieser Arbeit implementierten und getesteten Algorithmen zur Lösung des gegebenen nichtlinearen Optimierungsproblems mit nichtlinearen Nebenbedingungen vorgestellt. Die Auswertungen anhand einer Testfunktion zeigen sehr gute Ergebnisse. Da es sich bei der Testfunktion um eine reellwertige Funktion in zwei Veränderlichen handelt, sind diese Resultate zudem gut nachvollziehbar. Im Anschluss ist es uns gelungen ein Problem aus dem realen Verdichterprofildesign zu optimieren. Die Schwierigkeit bestand darin, das gegebene Mehrzieloptimierungsproblem in geeigneter Weise auf ein Problem mit einer reellwertigen Zielfunktion zu reduzieren. In diesem Zusammenhang wurden verschiedene Herangehensweisen aufgezeigt, um als Lösung der Ersatzmodell-Optimierung Verdichterprofile zu erhalten, deren wahre Funktionswerte nahe der Pareto-Front liegen.

Beim Vergleich der beiden implementierten und getesteten Verfahren zur Lösung des Problems mit nichtlinearen Nebenbedingungen zeigte sich kein eindeutiges Bild. In einigen Fällen konvergierte das Augmented-Lagrange-Verfahren eindeutig schneller, in anderen Fällen wies die Barriere-Methode eine geringere Iterationsanzahl auf. Zudem waren Unterschiede je nach Verwendung der unrestringierten Verfahren erkennbar. Bei diesen zeigte das Quasi-Newton-Verfahren im Vergleich das meist schnellste Konvergenzverhalten. Nur in wenigen Fällen konvergierte ein anderes Verfahren zur unrestringierten Minimierung in einer geringeren Anzahl an Iterationen. Dessen ungeachtet wies das Quasi-Newton-Verfahren hier ein akzeptables Verhalten auf. In diesem Zusammenhang sei jedoch darauf hingewiesen, dass die Verfahren aufgrund der unterschiedlichen Verfahrensvorschriften eventuell gegen verschiedene lokale Minima konvergieren. Ist man letztendlich an einem möglichst geringen Funktionswert interessiert, so kann die Untersuchung aller unrestringierten Verfahren ungeachtet der besseren Konvergenzeigenschaften des Quasi-Newton-Verfahrens hilfreich sein.

In den vorherigen Kapiteln wurden mit dem Augmented-Lagrange-Verfahren und der logarithmischen Barriere-Methode zwei unterschiedliche Verfahren für nichtlineare Optimierungsprobleme unter nichtlinearen Nebenbedingungen implementiert und untersucht. Zur Lösung eines Optimierungsproblems dieser Form existieren jedoch noch andere, im Rahmen dieser Arbeit nicht betrachtete Verfahren. Hier sei beispielsweise auf die weiteren, in Nocedal und Wright [36], Abschnitt 17.4, dargestellten Formulierungsmöglichkeiten des Augmented-Lagrange-Verfahrens für nichtlineare Nebenbedingungen verwiesen. Ferner existieren verschiedene, allgemein als Innere-Punkte-Methoden bezeichnete Herangehensweisen zur Lösung des Optimierungsproblems mit nichtlinearen Nebenbedingungen. Eine

Auswahl dieser Verfahren ist in Nocedal und Wright [36], Kapitel 19, verzeichnet. Somit besteht zusätzlich zu den im Rahmen dieser Arbeit implementierten Algorithmen die Möglichkeit, eines der eben genannten Verfahren für die Verdichteroptimierung zu verwenden oder schon implementierte Programmpakete zur nichtlinearen Optimierung mit nichtlinearen Nebenbedingungen zu nutzen (vgl. bspw. [36], Abschnitt 19.9).

Da die „Line Search“ des Armijo-Verfahrens teils viele Funktionsauswertungen nach sich zieht, könnte im Bereich der Schrittweitenberechnung noch Verbesserungspotential bestehen. Eine Erweiterungsmöglichkeit der bisher implementierten Verfahren stellt hier beispielsweise das Wolfe-Powell-Verfahren zur effizienten Schrittweitenberechnung dar (vgl. [35], S. 95 ff.). Im Rahmen dieser Arbeit haben wir uns für die Implementierung des Armijo-Verfahrens entschieden, da es effiziente Schrittweiten berechnet und im Vergleich zum Wolfe-Powell-Verfahren nur eine Funktionsauswertung pro Iteration benötigt. Während beim Armijo-Verfahren lediglich $\Phi(x^{(k)} + \sigma_j d^{(k)})$ in jeder Iteration ausgewertet wird (vgl. Verfahren 4.6), muss beim Wolfe-Powell-Verfahren zusätzlich der Gradient $\nabla \Phi(x^{(k)} + \sigma_j d^{(k)})$ berechnet werden. Hier wäre zu untersuchen, ob das Wolfe-Powell-Verfahren trotz einer erhöhten Anzahl an Funktionsauswertungen pro Iteration aufgrund durchschnittlich geringerer Iterationsanzahlen insgesamt weniger Auswertungen als das Armijo-Verfahren nach sich zieht. Die Anzahl der Funktionsauswertungen ist insbesondere im Hinblick auf eine mögliche Formoptimierung der wahren Funktion relevant (vgl. folgender Abschnitt und Herangehensweise in Abschnitt 1.1).

Die implementierten Algorithmen wurden im Rahmen dieser Arbeit zur Gradienten-basierten Optimierung von Ersatzmodellen eingesetzt. Mit Hilfe der implementierten Verfahren kann zukünftig die Formoptimierung der Verdichterprofile jedoch ebenso mit den wahren Funktionen und folglich ohne die Ersatzmodelle durchgeführt werden. Die hierfür benötigten Funktionsauswertungen, also beispielsweise im Fall von Abschnitt 5.2 die Werte des Totaldruckverhältnisses und Totaldruckverlustes in Abhängigkeit der fünf freien Variablen β_{ST} , β_{LE} , β_{TE} , y_{SS} und MA_{in} , berechnet der Strömungslöser TRACE (vgl. Abschnitt 2.2.2, S. 10). Die Ableitungen bezüglich der freien Variablen, auch als Formableitung (engl. *shape derivative*) bezeichnet, liefert sodann der adjungierte Strömungslöser adjointTRACE. Dieser löst die diskret adjungierten Reynolds-Gleichungen (vgl. [11]).

Die Formoptimierung der wahren Funktion ist im Vergleich zur Ersatzmodell-gestützten Optimierung selbstverständlich weitaus zeitintensiver, da jede Funktionsauswertung einen Durchlauf der Design-Prozesskette (vgl. Abschnitt 2.2.2) zur Folge hat und somit unter anderem die Lösung der RANS-Gleichung (2.1) in diskretisierter Form erfordert. Des Weiteren nimmt die Berechnung des Gradienten beziehungsweise der Formableitungen aufgrund der Lösung der diskret adjungierten Reynolds-Gleichungen nun wesentlich mehr Zeit in Anspruch.

Zur unmittelbaren Formoptimierung ohne Ersatzmodelle sind die im Rahmen dieser Arbeit implementierten Verfahren folglich nur dann sinnvoll einsetzbar, wenn die in Abschnitt 4.4.1 vorgeschlagene Parallelisierung der Line Search zur Schrittweitenberechnung im Armijo-Verfahren tatsächlich umgesetzt wird.

A Begriffserklärungen

A.1 Einführung in die Mehrzieloptimierung

Im Folgenden stellen wir die Prinzipien der Mehrzieloptimierung vor. Für weiterführende Literatur sei auf Ehrgott [9] verwiesen.

In der Mehrzieloptimierung werden Problemstellungen der Form

$$„\min_{x \in \mathcal{F}}“(f_1(x), \dots, f_q(x)). \quad (\text{A.1})$$

untersucht. Hierbei ist $\mathcal{F} \subseteq \mathbb{R}^n$ die zulässige Menge und $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$ sind für $i = 1, \dots, q$ die zu minimierenden Funktionen, wobei die Art der Minimierung noch zu klären ist. Eine eindeutige Form der Minimierung, wie es für Funktionen $f : \mathbb{R}^n \rightarrow \mathbb{R}$ der Fall ist, existiert hier aufgrund der fehlenden natürlichen Ordnungsstruktur des \mathbb{R}^q für $q > 1$ nicht.

Aus diesem Grund ist die Lösung einer Mehrzieloptimierungsaufgabe kein einzelner Punkt, sondern eine Menge an Kompromisslösungen, die sogenannte Pareto-Menge. Bevor wir diesen Begriff näher erläutern, definieren wir die Dominanz-Relation (vgl. [48]).

Definition A.1. Gegeben sei das Mehrzieloptimierungsproblem (A.1). Punkt $x \in \mathcal{F}$ dominiert $y \in \mathcal{F}$, falls

$$\begin{aligned} f_i(x) &\leq f_i(y) && \text{für alle } i \in \{1, \dots, q\} \text{ und} \\ f_j(x) &< f_j(y) && \text{für mindestens ein } j \in \{1, \dots, q\}. \end{aligned}$$

Zwei Punkte x und y können also in drei möglichen Varianten in Beziehung zueinander stehen: x dominiert y , x wird durch y dominiert oder x und y dominieren einander nicht.

Definition A.2. Ein Punkt $x \in \mathcal{F}$ heißt nichtdominiert genau dann, wenn kein Punkt $y \in \mathcal{F}$ existiert, welcher x dominiert. Der Punkt x heißt dann Pareto-Optimum. Eine Menge $X \subseteq \mathcal{F}$ heißt Pareto-Menge genau dann, wenn sie nur aus nichtdominierten Punkten besteht.

Der Pareto-Rang eines Punktes gibt die Anzahl der Punkte an, welche diesen Punkt dominieren.

Ein Punkt befindet sich somit in der Pareto-Menge, wenn die Minimierung eines Zielfunktionswerts nur noch unter gleichzeitiger Erhöhung eines anderen stattfinden kann (siehe Abb. A.1). Die Pareto-Menge wird aufgrund ihrer Form oft auch als Pareto-Front bezeichnet. Ist ein Punkt nicht in der Pareto-Menge enthalten, so kann dieser keinesfalls Lösung des Mehrzielproblems (A.1) sein, da es möglich ist, einen Zielfunktionswert weiter zu minimieren, ohne dabei einen anderen zu erhöhen. Ein Pareto-Optimum, also ein Element der Pareto-Menge, hat den geringsten Pareto-Rang 0.

Im Fall der Mehrzieloptimierung ist die Art der Minimierung nicht mehr eindeutig bestimmt. Eine Möglichkeit besteht in der Betrachtung einer gewichteten Summe der Zielfunktionen, sodass wir ein Optimierungsproblem der Form

$$\min_{x \in \mathcal{F}} \sum_{k=1}^q \lambda_k f_k(x)$$

mit Gewichtungsfaktoren $\lambda_k \in \mathbb{R}$ erhalten. Für weitere Herangehensweisen verweisen wir auf Ehrgott [9].

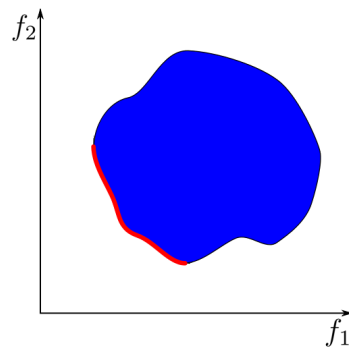


Abbildung A.1: Pareto-Menge (rot) einer Mehrzieloptimierung mit zwei Zielfunktionen und zweidimensionaler Wertemenge (blau) (Quelle: [49]).

A.2 Begriffe zur Parallelisierung

Im folgenden Abschnitt werden alle in Abschnitt 4.4.2 verwendeten Begriffe aus dem Gebiet der Parallelisierung kurz erläutert. Weitere Informationen zu diesem Thema finden sich beispielsweise in Precht et al. [50].

Parallelrechner, Knoten und Overhead

Bei Parallelrechnern handelt es sich um Rechner, die auf vielen Prozessoren gleichzeitig arbeiten. Die Grundbausteine eines Parallelrechners bilden die Knoten. Sie bestehen aus einem oder mehreren Hauptprozessoren und dazugehörigem gemeinsamen Speicher, auf den alle Prozessoren Zugriff haben (engl. *shared memory*). Zur Parallelisierung werden diese Knoten dann miteinander verknüpft. Hierbei reduziert sich die Rechenleistung des Parallelrechners im Vergleich zur Leistung aller einzelnen Knoten, da ein erheblicher Kommunikationsaufwand, der sogenannte Overhead, nötig ist. Somit ist neben der Rechenleistung der Knoten auch die Kommunikationsleistung von entscheidender Bedeutung für die Schnelligkeit eines Systems (vgl. [50]).

Prozessor, Prozessorkern und Mehrkernprozessor

Der Prozessor bildet die zentrale Recheneinheit des Computers und führt die Befehle aus. Das eigentliche Rechenwerk des Prozessors ist jedoch der Prozessorkern (engl. *core*). Besitzt ein Prozessor mehrere Kerne, so bezeichnet man ihn als Mehrkernprozessor. Hierzu zählt beispielsweise der in privaten Rechnern verwendete Dual-Core-Prozessor.

Lastverteilung und Skalierbarkeit

Die Lastverteilung (engl. *load balancing*) beschäftigt sich mit dem Problem der optimalen Aufteilung der Aufgaben auf die einzelnen Prozessoren. Dies ist sehr schwierig und selten ideal möglich. Es gibt immer wieder Prozessoren, die unbeschäftigt auf die Beendigung einer Teilaufgabe eines anderen Prozessors warten. Dies und die Kommunikation der parallelen Prozesse untereinander (vgl. Overhead) sind der Grund, weshalb eine Parallelisierung mit r Prozessoren selten einen Beschleunigungsfaktor von r zur Folge hat, sondern meist einen geringeren.

In diesem Zusammenhang sei der Begriff der Skalierbarkeit erwähnt. Bei idealer Skalierbarkeit wäre in obigem Beispiel ein Beschleunigungsfaktor von r der Fall. Spricht man von guter Skalierbarkeit, so befindet sich der Beschleunigungsfaktor nahe dem idealen. Bei steigender Anzahl an Prozessoren sinkt in der Regel die Skalierbarkeit, da häufig die Kommunikation gegenüber den reinen Rechenoperationen einen immer größeren Anteil der Gesamtausführungszeit eines Programms in Anspruch nimmt (vgl. [51]).

MPI, OpenMP und hybride Parallelisierung

OpenMP und MPI sind zwei Standards zur parallelen Programmierung, die auf unterschiedlichen Konzepten zur Kommunikation beruhen. OpenMP (Open Multi-Processing) benötigt ein System mit gemeinsamem Speicher (engl. *shared memory*), auf den alle Prozesse zugreifen können. Über diesen erfolgt auch der Datenaustausch. Beispielsweise kann es sich dabei um den gemeinsamen Speicher der (Mehrkern-)Prozessoren eines Knotens handeln (vgl. Seite 96). Im Kontrast dazu stehen Systeme mit verteiltem Speicher (engl. *distributed memory*), auf die der MPI-Standard (Message Passing Interface) anwendbar ist. Hier findet die Kommunikation der Prozesse mittels Nachrichten statt (siehe [52, 53]). Die simultane Verwendung von OpenMP und MPI wird als hybride Parallelisierung bezeichnet. In der Regel wird MPI hierbei zur groben Parallelisierung eingesetzt, während man mittels OpenMP die feinere Parallelisierung einzelner Programmschleifen vornimmt (siehe [51]).

B Implementierung der Verfahren

Alle Verfahren wurden in der Programmiersprache C programmiert. Die Auswertungen fanden unter dem Betriebssystem GNU/Linux, Version 2.6.32 mit der Distribution open-SUSE, Version 11.1 statt.

Im Rahmen dieser Arbeit wurde ein ausführbares Programm *Gradient.exe* erstellt. Die notwendigen Eingabedateien zur Wahl der Verfahren, Einstellung aller Parameter sowie Festlegung des Startpunktes und der Grenzen der freien Variablen x_i und Nebenbedingungen $h_j(x)$ müssen als Textdatei im Ordner *../Input* vor Beginn der Optimierung vorhanden sein. Hierbei handelt es sich um die Dateien *gradient.input*, *OptimizationParameters* sowie *limits.h.input*. Genauere Informationen sowie Beispiel-Eingabedateien finden sich in Abschnitt B.1.

Bei jeder Optimierung wird automatisch eine mit einem Zeitstempel versehene Ausgabe-datei im Ordner *../Output.GradientMethods* erzeugt. Diese enthält eine detaillierte Auflis-tung des Optimierungslaufes. Es werden die Iterationszahl, der verwendete Algorithmus, die aktuellen Variablen x_i , der Funktionswert $f(x)$ der zu minimierenden Funktion, die Nebenbedingungsverletzung $n(x)$, der Gradient $\nabla f(x)$, die Nebenbedingungen $h_j(x)$ und die Jacobi-Matrix $J_h(x)$ für jede Iteration des Armijo-Verfahrens, der unrestringierten Verfahren und der Augmented-Lagrange- respektive Barriere-Methode zeilenweise gespei-chert.

Das Interface, also die Schnittstelle zum Austausch von Daten zwischen dem im Rahmen dieser Arbeit implementierten Optimierer und den Ersatzmodellen wird in Abschnitt B.2 näher erläutert.

B.1 Eingabedateien

Im Folgenden werden die drei zur Optimierung notwendigen Eingabedateien vorgestellt. Das Rautenzeichen # signalisiert auskommentierte Zeilen.

Eingabedatei *gradient.input*

Die Datei *gradient.input* ist die grundlegende Eingabedatei zur Definition des zu optimie-renden Problems. Es werden die zu verwendenden Verfahren, die Art des Problems, die Form der Abbruchbedingung sowie alle Parameter festgelegt. Die folgende Eingabedatei orientiert sich an den in Kapitel 4 verwendeten Begriffen und mathematischen Symbolen. Sie stellt eine Beispieldatei zur Optimierung eines Problems der Form (P) (vgl. Kapitel 4, S. 35) mit Hilfe des Augmented-Lagrange-/Fletcher-Reeves-Verfahrens bei Anwendung der prozentualen Abbruchbedingung (vgl. Abschnitt 4.1, S. 36) dar.

```
#gradient.input
```

```
OPTIMIZATION_METHOD AUGMENTED_LAGRANGE
```

```
#OPTIMIZATION_METHOD BARRIER_METHOD
```

```
UNCONSTRAINED_OPTIMIZATION_METHOD FLETCHER_REEVES
```

```
#UNCONSTRAINED_OPTIMIZATION_METHOD POLAK_RIBIERE
```

```
#UNCONSTRAINED_OPTIMIZATION_METHOD GRADIENT_METHOD
```

```
#UNCONSTRAINED_OPTIMIZATION_METHOD QUASI_NEWTON
```

STEP_SIZE_METHOD ARMIJO

#number of constraints, cf. chapter 4, page 35

NUMBER_CONSTRAINTS m

#Information about the problem to minimize, cf. chapter 4, page 35

#FLAG_PROBLEM 1 for $\min f(x), x_i \in [a_i, b_i], h_j(x) \in [c_j, d_j]$

#FLAG_PROBLEM 2 for $\min f(x), x_i \in [a_i, b_i]$

FLAG_PROBLEM 1

#type of break condition for constrained method,

#cf. chapter 4.1, page 36 and page 37

#BREAK_CONDITION NORMAL

BREAK_CONDITION PERCENTAGE p_1 p_2 p_3 p_4

#tolerance for convergence test in constrained method,

#cf. chapter 4.1, page 36

TOLERANCE ε

#tolerance for convergence on zero gradient in unconstrained method,

#cf. chapter 4.3.1, page 40

TOL_GRAD ω

#tolerance for convergence on delta_x in unconstrained method,

#cf. chapter 4.3.1, page 40

TOL_X ω_x

#maximum number of iterations for constrained method

MAX_ITER max_1

#maximum number of iterations for unconstrained method

MAX_ITER_GRAD max_2

#maximum number of iterations for Armijo step size method

MAX_ITER_STEP max_3

#parameter in Armijo step size method, cf. chapter 4.3.5, page 46

DELTA_STEP δ

#lower boundary for σ_0 in Armijo, cf. chapter 4.3.5, page 46

SIGMA_SHORT 0.001

#parameter for step size reduction in Armijo method,

#cf. chapter 4.3.5, page 46

BETA_STEP β

Bemerkung B.1. Es sei darauf hingewiesen, dass die vorhandenen Optionen bei Festlegung des zu optimierenden Problems (vgl. FLAG_PROBLEM) keine Einschränkungen darstellen, da die oberen und unteren Grenzen der freien Variablen x_i und Nebenbedingungen $h_j(x)$ die Werte $-\infty$ respektive ∞ annehmen können.

Ist keine Nebenbedingung $h_j(x)$ vorhanden, so ist FLAG_PROBLEM 2 und $m = 0$ zu wählen. Die zwei möglichen Abbruchkriterien NORMAL respektive PERCENTAGE beziehen sich auf die beiden zur Verfügung stehenden Abbruchkriterien für das Augmented-Lagrange-Verfahren

beziehungsweise die Barriere-Methode. Bei der Wahl **NORMAL** wird die Differenz von aufeinanderfolgenden Funktionswerten und Iterierten betrachtet (vgl. Abschnitt 4.1, S. 36). Der zugehörige Parameter zur Festlegung der Abbruchtoleranz ist ε . Bei Wahl der Abbruchbedingung **PERCENTAGE** findet die prozentuale Abbruchbedingung Anwendung (vgl. Abschnitt 4.1, S. 37). Die zugehörigen Parameter p_1 bis p_4 werden hier direkt hinter **BREAK_CONDITION PERCENTAGE** eingetragen.

Ist ein bestimmtes Abbruchkriterium nicht erwünscht, wie etwa das Abbruchkriterium (4.7) (vgl. S. 40) für unrestringierte Verfahren basierend auf dem Abstand zweier Iterierter, so kann der entsprechende Parameter, in diesem Fall ω_x , Null gesetzt werden.

Bemerkung B.2. Viele der in der *gradient.input*-Datei angegebenen Auswahloptionen und Parameter können *online*, d.h. während der laufenden Optimierung angepasst und geändert werden. So kann etwa das Verfahren zur unrestringierten Minimierung in jeder inneren Iteration des Augmented-Lagrange-Verfahrens respektive der Barriere-Methode neu ausgewählt werden. Auch eine Anpassung der Parameter ω , ω_x , δ , **SIGMA_SHORT** und β sowie der maximalen Iterationsanzahl der unrestringierten Verfahren und des Armijo-Verfahrens sind möglich.

Eingabedatei *OptimizationParameters*

Mit Hilfe der Eingabedatei *OptimizationParameters* werden die Startwerte sowie die oberen und unteren Grenzen der freien Variablen x_i , $i \in \{1, \dots, n\}$, angegeben. Das folgende Beispiel zeigt die *OptimizationParameters*-Datei für die in Abschnitt 5.2.3 durchgeführte Auswertung. In der ersten Spalte ist der Variablenname angegeben. Darauf folgen die untere Grenze, der Startwert und die obere Grenze. Die letzte Spalte besitzt für unsere Optimierung keine Relevanz.

#OptimizationParameters

beta_ST	100	107.72	120	../../../../Slave\$/Profil/progen.input
beta_LE	120	137.94	150	../../../../Slave\$/Profil/progen.input
beta_TE	90	90.24	110	../../../../Slave\$/Profil/progen.input
YSS	0.1	0.30	0.3	../../../../Slave\$/Profil/progen.input
MA_in	0.7	0.77	0.78	../../../../Slave\$/Mises/ises.standard

Besitzt eine freie Variable keine untere beziehungsweise obere Grenze, so ist in diesem Fall **-INFINITY** respektive **INFINITY** an entsprechender Stelle einzutragen.

Eingabedatei *limits_h.input*

Die letzte Eingabedatei *limits_h.input* legt, falls vorhanden, die oberen und unteren Grenzen der Nebenbedingungen $h_j(x)$ fest. Die erste Spalte enthält die untere Grenze und die darauffolgende Spalte entsprechend die obere Grenze der jeweiligen Nebenbedingung. Zeilenweise werden nacheinander jeweils untere und obere Grenze in natürlicher Reihenfolge der Nebenbedingungen angegeben, mit $h_1(x)$ startend und $h_m(x)$ endend. Das folgende Beispiel zeigt die *limits_h.input*-Datei für die in Abschnitt 5.2.3 im ersten Unterabschnitt durchgeführte Auswertung.

```
#limits.h.input

#c_j      d_j
-INFINITY 0.04
-INFINITY 0.02
```

B.2 Interface zwischen Optimierer und Direct Gradient Enhanced Kriging-Verfahren

Der Datenaustausch zwischen dem im Rahmen dieser Arbeit implementierten Optimierer und dem Direct GEK erfolgt über Textdateien. Das Kriging-Verfahren berechnet die vom Optimierer benötigten Funktionswerte des Ersatzmodells, des approximierten Zielfunktionsgradienten, der Nebenbedingungen und deren Ableitungen für die aktuelle Variable x .

Alle zum Datenaustausch notwendigen Textdateien befinden sich in dem Ordner *../Output.GradientMethods*. Der Optimierer speichert vor Beginn der Funktionsauswertung durch das Direct GEK die aktuelle Variable x als Spalte der Länge n in der Datei *freeVar.txt*, sodass das Direct GEK auf diese Zugriff hat. Im Anschluss wird von Seiten des Optimierers der Konsolenbefehl

./UDP.exe ../Input/UserDefinedProcessChain

abgesetzt. In diesem Fall soll das Direct GEK den Wert des Ersatzmodells, den Gradienten, die Nebenbedingungen und deren Ableitungen berechnen und ausgeben. Soll nur eine der vier zuvor genannten Funktionen ausgewertet werden, so muss die Eingabedatei des Konsolenbefehls entsprechend abgeändert werden. Wird einzig der Funktionswert des Ersatzmodells benötigt, so muss *../Input/UserDefinedProcessChain_f* übergeben werden. Zur Auswertung des Gradienten, der Nebenbedingungen oder deren Ableitungen ändert sich der Zusatz entsprechend zu *_df*, *_h* respektive *_Jach*.

Der Befehl *./UDP.exe* mit entsprechender Datei als Übergabeparameter signalisiert dem Direct GEK, welche Funktion ausgewertet werden soll. Das Verfahren erzeugt daraufhin entsprechende Textdateien, welche die geforderten Funktionswerte enthalten. Ist der approximierte Zielfunktionswert beziehungsweise Ersatzmodellwert gefordert, so wird die Ausgabedatei *UDP_f.output* erzeugt. Diese enthält eine Zahl. Bei Berechnung des Gradienten wird die Ausgabedatei *UDP_df.output* erzeugt, welche den geforderten Gradienten als Spalte der Länge n enthält. Die Nebenbedingung $h(x)$ wird als Spalte der Länge m in *UDP_h.output* gespeichert. Sind schließlich die Ableitungen der Nebenbedingung gefordert, so werden diese in Form der Jacobi-Matrix gespeichert. Die Datei *UDP_Jach.output* enthält entsprechend m Zeilen und n Spalten.

Nach Berechnung der geforderten Funktionswerte durch das Direct GEK liest der Optimierer die entsprechenden Ausgabedateien aus und führt die Optimierung mit Hilfe dieser Werte fort.

Literatur

- [1] SILLER, Ulrich; VOSS, Christian ; NICKE, Eberhard: Automated Multidisciplinary Optimization of a Transonic Axial Compressor. In: *47th AIAA Aerospace Sciences Meeting Including The New Horizons Forum and Aerospace Exposition*. Orlando, Florida, Januar 2009
- [2] KROLL, Norbert; BECKER, Klaus; RIEGER, Herbert ; THIELE, Frank: Ongoing activities in flow simulation and shape optimization within the German MEGADESIGN project, 25th International Congress of the Aeronautical Sciences Proceedings, 2006
- [3] *Wikipedia-Homepage: Verdichter*. <http://de.wikipedia.org/wiki/Verdichter>, Abruf: 14. Januar 2012
- [4] *Ratgeber Druckluft*. <http://www.drucklufttechnik.de/>, Abruf: 14. Januar 2012
- [5] SIEKMANN, Helmut E.; THAMSEN, Paul U.: *Strömungslehre Grundlagen*. 2. Auflage. Springer Lehrbuch, 2008
- [6] *Wikipedia-Homepage: Strahltriebwerk*. <http://de.wikipedia.org/wiki/Strahltriebwerk>, Abruf: 14. Januar 2012
- [7] *DLR-Homepage: AutoOpti*. http://www.dlr.de/at/desktopdefault.aspx/tabid-1529/2164_read-4771/, Abruf: 14. Januar 2012
- [8] VOSS, Christian: *AutoOpti: Overview*. März 2010. – Vortragsfolien
- [9] EHRGOTT, Matthias: *Multicriteria Optimization*. 2. Auflage. Springer Verlag, 2005
- [10] MACKAY, David J. C.: *Bayesian Methods for Adaptive Models*, California Institute of Technology, Diss., 1992
- [11] BACKHAUS, Jan; AULICH, Marcel; FREY, Christian; LENGYEL, Timea ; VOSS, Christian: *Gradient enhanced surrogate models based on adjoint CFD methods for the design of a counter rotating turbofan*, Proceedings of ADME Turbo Expo 2012, 2012
- [12] WEBER, Anton: 3D Structured Grids for Multistage Turbomachinery Applications based on G3DMESH. In: *DLR IB-325-05-04* (2004)
- [13] KERSKEN, Hans-Peter; FREY, Christian; VOIGT, Christian ; ASHCROFT, Graham: *Time-linearized and Time-accurate 3D RANS Methods for Aeroelastic Analysis in Turbomachinery*, Proceedings of ASME Turbo Expo 2010: Power for Land, Sea and Air, 2010
- [14] WILCOX, David C.: Re-assessment of the scale-determining equation for advanced turbulence models. In: *AIAA Journal*, vol 26, no. 11 (1988), S. 1299–1310
- [15] BAUMANN, Wolfgang W.; BUNGE, Ulf; SCHATZ, Markus ; THIELE, Frank: Version: 4. Auflage, 2003. http://www.cfd.tu-berlin.de/Lehre/fvm_skript/, Abruf: 24. Mai 2012. Technische Universität Berlin
- [16] ZÖLLNER, Melven: *Parallele, iterative Lösung dünnbesetzter linearer Gleichungssysteme aus der Strömungstechnik*, Rheinisch-Westfälische Technische Hochschule Aachen, Bachelorarbeit, 2011

- [17] Homepage INTES (Ingenieurgesellschaft für technische Software): Software PERMAS. <http://www.intes.de/>, Abruf: 24. Mai 2012
- [18] CalculiX-Homepage. <http://www.calculix.de/>, Abruf: 24. Mai 2012
- [19] STREIT, Ulrich: *Online-Skript zur Übung „Werkzeuge zur numerischen Modellierung“*. http://ifgivor.uni-muenster.de/vorlesungen/Num_Modellierung/, Abruf: 22. November 2011
- [20] KISSAUER, Stefan: *Hausarbeit zum Thema: Geostatistik II - Kriging Interpolation*. http://www.geogr.uni-jena.de/~c5hema/gis_ws04/ha/kissauer.pdf, Abruf: 21. November 2011
- [21] WU, Zongmin: *Die Kriging-Methode zur Lösung mehrdimensionaler Interpolationsprobleme*, Georg-August-Universität zu Göttingen, Diss., 1986
- [22] SHIBLI, Syed Abdul R.: *Geostatistics FAQ - Frequently Asked Questions*. http://www.ai-geostats.org/pub/AI_GEOSTATS/AI_GEOSTATSFAQ/FAQ_Geostatistics_01.pdf, Abruf: 21. November 2011
- [23] KRIGE, Daniel G.: *A statistical approach to some mine valuation and allied problems on the Witwatersrand*, University of the Witwatersrand, Masterarbeit, 1951
- [24] MATHERON, Georges: Principles of geostatistics. In: *Economic Geology*, vol 58 (1963), S. 1246–1266
- [25] *englische Wikipedia-Homepage: Variogram*. <http://en.wikipedia.org/wiki/Variogram>, Abruf: 23. November 2011
- [26] ISAACS, Edward H.; SRIVASTAVA, Mohan: *An introduction to applied geostatistics*. Oxford University Press, Inc., 1989
- [27] SCHULZ, Eric: *Anwendung von geostatistischen Methoden in ArcGIS zur Regionalisierung von Grundwasseranalysen am Beispiel eines Benzol-Grundwasserschadens*, Hochschule für Technik und Wirtschaft Dresden, Diplomarbeit, 2009
- [28] HATTERMANN, Fred: *Folien zur Vorlesung „Einführung in die Geostatistik“*. <http://www.pik-potsdam.de/~fred/geostatistik/>, Abruf: 24. November 2011
- [29] CHILÈS, Jean-Paul; DELFINER, Pierre: *Geostatistics: Modeling Spatial Uncertainty*. Wiley Series in Probability and Statistics, 1999
- [30] JANG, Jyh-Shing R.: *Invertieren einer Blockmatrix*. <http://www.cs.nthu.edu.tw/~jang/book/addenda/matinv/matinv/>, Abruf: 20. Januar 2012
- [31] ARCGIS RESOURCE CENTER: *Funktionsweise des Werkzeugs "Kriging"*. <http://help.arcgis.com/de/arcgisdesktop/10.0/help/index.html#//00q900000001t000000.htm>, Abruf: 5. Dezember 2011
- [32] YAMAZAKI, Wataru; RUMPFKEIL, Markus P. ; MAVRIPLIS, Dimitri J.: *Design Optimization Utilizing Gradient/Hessian Enhanced Surrogate Models*, 28th AIAA Applied Aerodynamics Conference, 2010

-
- [33] FIACCO, Anthony V.; MCCORMICK, Garth P.: *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*. Wiley, New York, 1968
 - [34] FIACCO, Anthony V.; MCCORMICK, Garth P.: *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*. SIAM, Philadelphia, 1990 (Classics in Applied Mathematics)
 - [35] ALT, Walter: *Nichtlineare Optimierung. Eine Einführung in Theorie, Verfahren und Anwendungen*. 2. Auflage. Vieweg+Teubner Verlag, 2011
 - [36] NOCEDAL, Jorge; WRIGHT, Stephen J.: *Numerical Optimization*. Second Edition. Springer, 2006 (Springer Series in Operations Research)
 - [37] FLETCHER, Roger; REEVES, Colin M.: Function minimization by conjugate gradients. In: *Computer Journal*, vol 7 (1964), S. 149–154
 - [38] HARZHEIM, Lothar: *Strukturoptimierung, Grundlagen und Anwendungen*. Verlag Harri Deutsch, 2008
 - [39] POWELL, Michael J. D.: Restart procedures for the conjugate gradient method. In: *Mathematical Programming*, vol 12 (1977), S. 241–254
 - [40] SPELLUCCI, Peter: *Numerische Verfahren der nichtlinearen Optimierung*. Birkhäuser Verlag, Basel, 1993
 - [41] POLAK, Elijah; RIBIÈRE, Gerard: Note sur la convergence de méthodes de directions conjuguées. In: *Revue française d’informatique et de recherche opérationnelle, série rouge*, vol 3 (1969), S. 35–43
 - [42] PRESS, William H.; TEUKOLSKY, Saul A.; VETTERLING, William T. ; FIANNERY, Brian P.: *Numerical Recipes in C, The Art of Scientific Computing*. Second Edition. Cambridge University Press, 1992
 - [43] POLAK, Elijah: *Computational Methods in Optimization*. Academic Press, New York, 1971 (Mathematics in science and engineering)
 - [44] GOLDSTEIN, Allen A.: On Steepest Descent. In: *SIAM Journal on Control and Optimization*, vol 3 (1965), S. 147–151
 - [45] ARMIJO, Larry: Minimization of Functions Having Lipschitz Continuous First Partial Derivatives. In: *Pacific Journal of Mathematics*, vol. 16 (1966), S. 1–3
 - [46] BASERMANN, Achim; KERSKEN, Hans-Peter; SCHREIBER, Andreas; GERHOLD, Thomas; JÄGERSKÜPPER, Jens; KROLL, Norbert; BACKHAUS, Jan; KÜGELER, Edmund; ALRUTZ, Thomas; SIMMENDINGER, Christian; FELDHOFF, Kim; KRZIKALLA, Olaf; MÜLLER-PFEFFERKORN, Ralph; PUETZ, Mathias; AUMANN, Petra; KNOBLOCH, Olaf; HUNGER, Jörg ; ZSCHERP, Carsten: *HICFD – Highly Efficient Implementation of CFD Codes for HPC Many-Core Architectures*, Proceedings of an International Conference on Competence in High Performance Computing (2010), 2012, S. 1–14
 - [47] DRELA, Mark; YOUNGREN, Harold: *A User’s Guide to MISES 2.63*. Version: 2008. <http://web.mit.edu/drela/Public/web/mises/mises.pdf>, Abruf: 13. März 2012. Aerospace Computational Design Laboratory, MIT

- [48] STÖCKER, Martin: *Untersuchung von Optimierungsverfahren für rechenzeit-aufwändige technische Anwendungen in der Motorenentwicklung*, Technischen Universität Chemnitz, Diplomarbeit, 2007
- [49] *Wikipedia: Pareto-Optimum*. <http://de.wikipedia.org/wiki/Pareto-Optimum>, Abruf: 10. April 2012
- [50] PRECHT, Manfred; MEIER, Nikolaus ; TREMEL, Dieter: *EDV-Grundwissen: Eine Einführung in Theorie und Praxis der modernen EDV*. 7., aktualisierte Auflage. Addison-Wesley Verlag, 2004
- [51] RECHEN- UND KOMMUNIKATIONSZENTRUM RWTH AACHEN: *Parallelisierung*. <http://www.rz.rwth-aachen.de/aw/cms/rz/Themen/hochleistungsrechnen/programmierung/~pti/parallelisierung/?lang=de>, Abruf: 16. März 2012
- [52] MESSAGE PASSING INTERFACE FORUM: MPI: A Message-Passing Interface standard. In: *The International Journal of Supercomputer Applications and High Performance Computing*, vol 8 (1994), Nr. 3/4, S. 159–416
- [53] MESSAGE PASSING INTERFACE FORUM: *MPI-2: Extensions to the Message-Passing Interface*. Version: 1997. <http://www.mpi-forum.org/docs/mpi-20-html/mpi2-report.html>, Abruf: 11. März 2012. University of Tennessee, Knoxville, Tennessee

Abbildungsverzeichnis

1.1	Verdichterschaukeln	1
1.2	Skizze eines Verdichterschaukelprofils	1
2.1	Axialverdichter	5
2.2	Strahltriebwerk	6
2.3	Programmaufbau von AutoOpti	7
2.4	Gitter einer Verdichterstufe	10
2.5	FE-Strukturanalyse einer Verdichterschaukel	11
3.1	Polygon-Methode	14
3.2	Inverse Distanzgewichtung	14
3.3	Nachteil der Inversen Distanzgewichtung	15
3.4	Beispiel für Stützstellen und Ersatzmodell	17
3.5	Beispiel eines Semivariogramms	18
3.6	Verschiedene Semivariogrammfunktionen im Vergleich	22
5.1	Testfunktion f	52
5.2	Ersatzmodell der Testfunktion	52
5.3	Niveaumenge des Ersatzmodells	53
5.4	Konvergenzverlauf des Augmented-Lagrange-Verfahrens und der Barriere-Methode bei einem Minimum	54
5.5	Wege der Augmented-Lagrange-Iterierten bei einem Minimum	55
5.6	Konvergenz des Augmented-Lagrange-Verfahrens mit verschiedenen Werten für SIGMA.SHORT	56
5.7	Niveaumenge des Ersatzmodells mit Startpunkten	57
5.8	Konvergenzverlauf des Augmented-Lagrange-Verfahrens und der Barriere-Methode zum Startpunkt (0.5,0.9)	57
5.9	Konvergenzverlauf des Augmented-Lagrange-Verfahrens und der Barriere-Methode zu den Startpunkten (0.05,0.05) und (0.35,0.5)	59
5.10	Niveauflächen der Unsicherheit und des Ersatzmodells bei Unsicherheitsrestriktion von 0.7	60
5.11	Konvergenzverlauf des Augmented-Lagrange-Verfahrens und der Barriere-Methode bei Restriktion der Unsicherheit durch 0.7	61
5.12	Weg der Iterierten beim Augmented-Lagrange-Verfahren mit Restriktion der Unsicherheit durch 0.7	61
5.13	Weg der Iterierten beim Augmented-Lagrange-Verfahren mit Restriktion der Unsicherheit durch 0.8	63
5.14	Konvergenzverlauf des Augmented-Lagrange-Verfahrens und der Barriere-Methode bei Restriktion der Unsicherheit durch 0.8	64
5.15	Niveaufläche der Unsicherheit und des Ersatzmodells bei Restriktion durch 1.0	65
5.16	Konvergenzverlauf des Augmented-Lagrange-Verfahrens und der Barriere-Methode bei Restriktion der Unsicherheit durch 1.0	65
5.17	Niveaufläche der Unsicherheit und des Ersatzmodells bei Restriktion durch 0.95	67
5.18	Konvergenzverlauf des Augmented-Lagrange-Verfahrens und der Barriere-Methode bei Restriktion der Unsicherheit durch 0.95	67
5.19	Skizze eines Verdichterprofils mit geometrischen Variablen	70
5.20	Datenbasis des Ersatzmodells mit Pareto-Menge	71

5.21	Approximation und wahre Zielfunktionswerte bei Gewichtung mit -0.628 und 27.027	73
5.22	Konvergenzverlauf des Augmented-Lagrange-Verfahrens und der Barriere-Methode bei Gewichtung mit -0.628 und 27.027	74
5.23	Approximation und wahre Zielfunktionswerte bei Gewichtung mit -1.046 und 9.009	75
5.24	Konvergenzverlauf des Augmented-Lagrange-Verfahrens und der Barriere-Methode bei Gewichtung mit -1.046 und 9.009	76
5.25	Approximation und wahre Zielfunktionswerte bei Gewichtung mit -1.245 und 0.452	78
5.26	Konvergenzverlauf des Augmented-Lagrange-Verfahrens und der Barriere-Methode bei Gewichtung mit -1.245 und 0.452	79
5.27	Approximation und wahre Zielfunktionswerte bei Restriktion $p_{\text{TPL}} \leq 0.04$.	80
5.28	Approximation und wahre Zielfunktionswerte bei Restriktion $p_{\text{TPL}} \leq 0.03$.	82
5.29	Approximation und wahre Zielfunktionswerte bei Restriktion $p_{\text{TPL}} \leq 0.025$.	83
5.30	Approximation der Zielfunktionswerte bei Restriktion $p_{\text{TPL}} \leq 0.0325$	84
5.31	Approximation und wahre Werte der Zielfunktionswerte bei Restriktion $h_1(x) \leq 0.04$ und $h_2(x) \leq 0.02$	86
5.32	Norm des Differenzvektors bei unterschiedlichen Restriktionen	88
5.33	Approximation und wahre Zielfunktionswerte bei verschiedenen Restriktionen der Unsicherheit	91
A.1	Pareto-Menge	96